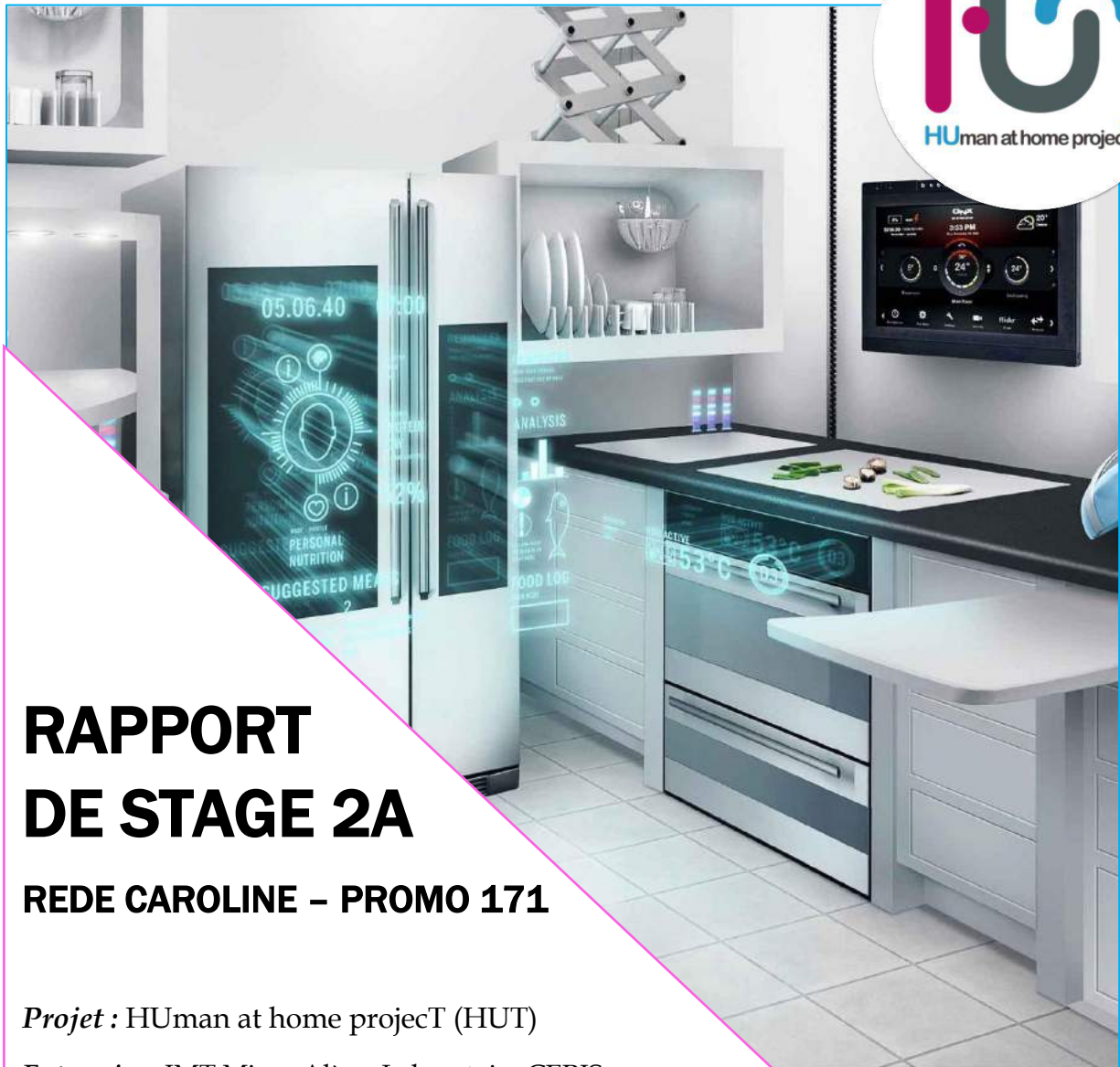


# STAGE DE RECHERCHE EN INFORMATIQUE SUR LES MAISONS CONNECTEES (SMART HOMES)



## RAPPORT DE STAGE 2A

**REDE CAROLINE – PROMO 171**

*Projet :* HUman at home project (HUT)

*Entreprise :* IMT Mines Alès – Laboratoire CERIS

*Durée :* 13 semaines – du 01 Mai au 29 Juillet 2022

Tuteur de stage : VAUTTIER Sylvain

Autres collaborateurs : PEREZ Quentin, URTADO Christelle



# Remerciements

Tout d'abord, j'aimerais remercier mon école, *IMT Mines Alès* et plus particulièrement mon professeur, **M. Sylvain VAUTTIER**, mon tuteur lors de ce stage, qui m'a offert cette opportunité de travailler sur un projet de recherche dans mon domaine de spécialisation. Cela m'a permis de découvrir de manière plus approfondie le milieu de la recherche, notamment sur un sujet qui m'inspire beaucoup pour mon futur parcours professionnel qui est la maison connectée de demain, autrement appelée *smart home*.

Ensuite, je souhaiterais remercier les autres personnes encadrantes qui dirigent ce projet conjointement, ma professeure **Mme Christelle URTADO**, ainsi que le doctorant **M. Quentin PEREZ**. Tous trois ont été de bons conseils pour nous orienter tout au long de ce stage. Ils nous ont aussi offert la possibilité de participer à une conférence en ligne sur le monde de la recherche, ceci afin d'en apprendre davantage sur ce domaine et une éventuelle poursuite d'étude, ce qui fut très enrichissant.

Enfin, je voudrais remercier mon camarade de classe et co-stagiaire **M. Aberrahmane BAROUALI** qui a travaillé avec moi tout au long de ces treize semaines, avec qui j'ai pu beaucoup échanger pour faire évoluer nos réflexions et nos idées sur le projet, et répondre aux objectifs attendus.

# Table des matières

<b>1</b>	<b>Introduction .....</b>	<b>8</b>
1.1	Cadre du stage .....	8
1.2	Présentation du projet HUT .....	8
1.3	Présentation de l'école et des différentes parties prenantes .....	9
1.4	Objectifs du stage .....	10
1.5	Démarche établie.....	10
<b>2</b>	<b>Réalisations de ce stage.....</b>	<b>11</b>
2.1	Introduction .....	11
2.2	Résumé du travail préliminaire .....	11
2.3	Analyse des données.....	12
2.3.1	Présentation du jeu de données .....	12
2.3.2	Objectifs et démarche.....	13
2.3.3	Analyse graphique des données.....	13
2.3.4	Analyse des données sous forme d'activités .....	16
2.4	Etablissement de scénarios .....	25
2.4.1	Première modélisation dite « simple ».....	26
2.4.2	Deuxième modélisation à 3 chaînes .....	33
<b>3</b>	<b>Conclusion.....</b>	<b>35</b>
3.1	Retour d'expérience .....	35
3.2	Suite du projet .....	35
	<b>Annexes .....</b>	<b>36</b>
	Description des rôles des différentes parties prenantes du projet .....	36
	Plan de l'appartement du jeu de données.....	37
	Tableau récapitulatif des capteurs et de leurs caractéristiques importantes .....	38
	Tableau des observations graphiques menées pour évaluer la présence dans l'appartement .....	40
	Résultats obtenus pour l'analyse de l'activité « regarder la télé » .....	48
	Résultats obtenus pour l'analyse de l'activité « dormir dans le lit » .....	50
	Résultats obtenus pour l'analyse de la corrélation entre deux activités : « regarder la télé » et « dormir dans le lit » .....	53
	Résultats obtenus pour l'analyse de l'activité « faire un repas » .....	54
	Résultats obtenus pour l'analyse de l'activité de la salle de bain.....	61
	Résultats obtenus pour l'analyse de l'apparition de la lettre suivante .....	65
	Résultats obtenus pour l'analyse de l'appartenance d'activités à un bloc .....	66
	Résultats obtenus pour l'analyse de l'apparition de la lettre suivante sans les capteurs de mouvement.....	67

<b>Résultats obtenus pour l'analyse de l'apparition de la lettre suivante avec un changement d'activité .....</b>	<b>67</b>
<b>Résultats obtenus pour l'analyse d'occurrence de séquences de caractères .....</b>	<b>68</b>
<b>Résultats obtenus pour l'analyse d'occurrence de séquences de caractères sans les capteurs de mouvement.....</b>	<b>69</b>
<b>Bibliographie et sources.....</b>	<b>71</b>
<b>Glossaire.....</b>	<b>71</b>
<b>Synthèse .....</b>	<b>72</b>
<b>Abstract.....</b>	<b>72</b>

## Table des illustrations

Figure 1 : Logo du projet HUT .....	8
Figure 2 : Logo de IMT Mines Alès .....	9
Figure 3 : Partenaires du projet HUT .....	10
Figure 4 : Visualisation des plages de l'activité "regarder la télé" pour l'ensemble des données .....	17
Figure 5 : Visualisation de la fréquence selon le début et la fin de l'activité "regarder la télé" .....	17
Figure 6 : Table des fréquences et durées moyennes pour l'activité "regarder la télé" .....	18
Figure 7 : Calendrier des fréquences et durées moyennes pour l'activité "regarder la télé" ..	19
Figure 8 : Plan de l'appartement du jeu de données étudié avec disposition des capteurs ...	37
Figure 9 : [image 1] Graphique de la porte d'entrée et mouvement dans l'entrée (données à la minute sur une semaine) .....	45
Figure 10 : [image 2] Graphique des capteurs de mouvement (données à la minute sur une journée) .....	45
Figure 11 : [image 5] Graphique des capteurs de mouvement et porte d'entrée (données toutes les 15 sec sur la journée) .....	46
Figure 12 : [image 9] Graphique des capteurs de mouvement livingroom / bathroom / couloir (données toutes les 15 sec sur la semaine) .....	46
Figure 13 : [image 10] Graphique des capteurs de mouvement livingroom / kitchen / couloir (données toutes les 15 sec sur la semaine) .....	46
Figure 14 : [image 11] Graphique des capteurs de mouvement livingroom / bedroom / couloir (données toutes les 15 sec sur la semaine) .....	47
Figure 15 : [image 12] Graphique des capteurs de mouvement bathroom / kitchen / couloir (données toutes les 15 sec sur la semaine) .....	47
Figure 16 : [image 13] Graphique des capteurs de mouvement bathroom / bedroom et pression du lit (données toutes les minutes sur la semaine) .....	47
Figure 17 : Visualisation des plages de l'activité "regardé la télé" pour l'ensemble des données .....	48
Figure 18 : Visualisation de la fréquence selon le début et la fin de l'activité "regarder la télé" .....	48
Figure 19 : Visualisation des plages pour l'activité "dormir dans le lit" .....	50
Figure 20 : Visualisation de la fréquence d'appartenance de l'activité à un créneau selon le début de la fin de l'activité "dormir dans le lit", l'indice 0 du graphe correspondant à 15:00 ..	50
Figure 21 : Visualisation des fréquences de l'activité "faire un repas" par la combinaison de 3 capteurs (frigo, machine à café, sandwichmaker), en considérant les données du frigo de 1 sec .....	54
Figure 22 : Visualisation de la fréquence de l'activité de la machine à café correspondant au début de l'activité .....	55
Figure 23 : Visualisation de la fréquence de l'activité du lave-vaisselle correspondant au début de l'activité .....	57
Figure 24 : Visualisation de la fréquence de l'activité correspondant au début de l'activité ...	59
Figure 25 : Visualisation de la fréquence de l'activité correspondant au début de l'activité ...	61
Figure 26 : Visualisation de la fréquence de l'activité correspondant au début de l'activité ...	63

## Table des tableaux

Tableau 1 : Conclusions des observations graphiques menées .....	15
Tableau 2 : Table des plages des deux activités "regarder la télé" et "dormir dans le lit" qui se suivent à 20 minutes .....	22
Tableau 3 : Tableau récapitulatif des capteurs, leur référence, leur type, les plages de données relevées, ainsi que leurs états.....	40
Tableau 4 : Tableau des analyses menées à partir de visualisation des données uniquement pour évaluer la présence dans l'appartement .....	45
Tableau 5 : Table des fréquences et durées moyennes pour l'activité "regardé la télé" .....	49
Tableau 6 : Calendrier des fréquences et durées moyennes pour l'activité "regardé la télé" .....	50
Tableau 7 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage.....	51
Tableau 8 : Calendrier des moyennes de durées et des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage ....	52
Tableau 9 : Table des plages des deux activités "regarder la télé" et "dormir dans le lit" qui se suivent à 20 minutes .....	53
Tableau 10 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage.....	55
Tableau 11 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage .....	56
Tableau 12 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage.....	57
Tableau 13 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage .....	58
Tableau 14 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage.....	59
Tableau 15 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage .....	60
Tableau 16 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage .....	61
Tableau 17 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage .....	62
Tableau 18 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage .....	63
Tableau 19 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage .....	64
Tableau 20 : Table des occurrences d'apparition des termes dans la chaîne .....	65
Tableau 21 : Table des occurrences d'apparition des termes dans la chaîne (doublons de blocs traités).....	65
Tableau 22 : Table des probabilités d'apparition de la lettre suivante (colonne) pour une lettre choisie (ligne) .....	65
Tableau 23 : Table des probabilités d'apparition de la lettre suivante (colonne) par rapport à une lettre choisie (ligne), considérant toutes les lettres d'un même bloc.....	66
Tableau 24 : Table des probabilités d'appartenance d'une lettre (colonne) au sein d'un même bloc qu'une lettre choisie (ligne) .....	66
Tableau 25 : Table des occurrences d'apparition des termes dans la chaîne sans capteurs de mouvement.....	67

Tableau 26 : Table des probabilités d'apparition de la lettre suivante (colonne), selon une lettre choisie (ligne) sans les capteurs de mouvement.....	67
Tableau 27 : Table des occurrences des termes dans la chaîne sans les capteurs de mouvement et pour la réunion de capteurs sous l'activité « faire un repas » (q) .....	67
Tableau 28 : Table des probabilités d'apparition de la lettre suivante (colonne) selon une lettre choisie (ligne) sans les capteurs de mouvement et pour la réunion de capteurs sous l'activité « faire un repas » (q).....	68
Tableau 29 : Table des probabilités d'apparition de séquences de caractères selon la longueur de la séquence N.B. : en surligné, les expressions qui ne paraissent pas sensées, a priori .....	69
Tableau 30 : Table des probabilités d'apparition de séquences de caractères selon la longueur de la séquence sans les capteurs de mouvement .....	70

# 1 Introduction

## 1.1 Cadre du stage

Dans le cadre de ma formation à l'IMT Mines Alès, j'ai eu l'opportunité d'effectuer mon stage de deuxième année en tant qu'assistante ingénieure au sein même de mon école, et plus particulièrement du laboratoire informatique de celui-ci, le *Laboratoire CERIS*. Ce stage s'est déroulé pour une durée de 13 semaines, du 01 Mai au 29 Juillet 2022, à Alès (30), sur le site de recherche de l'école communément appelé *Croupillac*.

Pour ce stage, outre les encadrants, j'ai collaboré tout du long avec un de mes camarades de classe. Nous avons alors travaillé ensemble pendant ce stage pour faire évoluer au mieux ce projet et répondre aux objectifs, en se partageant les différentes tâches mais aussi nos réflexions et idées sur le sujet. Ce rapport ne présentera que les études que j'ai réalisées et non celles de mon camarade, qui consignera les siennes dans son rapport. Le travail s'étant fait conjointement, les bases des modélisations employées peuvent être similaires, mais pour différentes analyses menées avec un jeu de données différent (un échantillon du jeu de données contre la totalité de celui-ci par exemple).

## 1.2 Présentation du projet HUT

HUT, qui signifie HUman at home projecT est un projet interdisciplinaire qui réunit plusieurs laboratoires, entreprises et institutions. Il a pour objectif d'étudier l'impact des effets des nouvelles technologies et des objets connectés sur les comportements et le bien-être des utilisateurs. A cet effet, un appartement observatoire équipé d'un environnement domotique et d'un ensemble de capteurs a été mis en place et occupé à trois reprises par des binômes volontaires, chacun sur une durée de 1 an. Les données prélevées par ces appareils sont analysées pour pouvoir étudier les comportements des utilisateurs et leurs évolutions, ainsi que leur usage du numérique dans leur quotidien et l'utilité de tels environnements.



Figure 1 : Logo du projet HUT

Il s'agit alors d'un projet orienté sciences humaines, puisqu'on s'intéresse aux comportements des utilisateurs dans un environnement connecté, basé sur des concepts scientifiques de par l'analyse de données importantes et le développement d'outils pour ces environnements futurs. Les possibles applications de ce type de projet sont nombreuses, englobées sous l'intitulé de *smart home*, telles que le l'enclenchement et l'automatisation d'appareils, le contrôle de la consommation énergétique ou encore l'aide à domicile dans la prévention aux accidents.



### 1.3 Présentation de l'école et des différentes parties prenantes

Le projet HUT est porté par différentes organisations telles que des industriels, des institutions et des laboratoires de recherche. Dans le cadre de mon stage, j'ai travaillé au sein de l'IMT Mines Alès et plus précisément du laboratoire informatique associé CERIS.

IMT Mines Alès fait partie du groupe *Institut Mines Télécom (IMT)*, qui est un établissement public scientifique, culturel et professionnel placé sous la tutelle principale des ministres chargés de l'industrie et du numérique. Il s'agit d'un grand groupe d'écoles d'ingénieurs en France, se composant de 11 écoles publiques répartis sur le territoire et formant 13 500 ingénieurs et docteurs. Cette organisation emploie 4 500 personnes, possède un budget annuel de 400M€ et comporte 2 instituts Carnot, 35 chaires industrielles, produit 2 100 publications de rang A, 60 brevets et réalise 110M€ de recherche contractuelle.



Figure 2 : Logo de IMT Mines Alès

Parmi cette organisation, IMT Mines Alès a été créée en 1843, implantée à Alès dans la région d'Occitanie, où elle dispose de deux campus. Elle comporte aujourd'hui 1 400 élèves, dont 250 internationaux, et 400 personnels. Cette école forme des élèves ingénieurs généralistes, des ingénieurs de spécialité (par apprentissage), ainsi que des doctorants et des élèves de masters. Elle se compose aussi de trois centres de recherches de haut niveau scientifique et technologique dans les domaines des matériaux et du génie civil (C2MA), de l'environnement et des risques (CREER), ainsi que de l'intelligence artificielle et du génie industriel et numérique (CERIS). Ces organisations regroupent 93 enseignants-chercheurs ainsi que 104 doctorants et post-doctorants. L'école fut aussi la première à créer un incubateur en 1984, créant ainsi 200 entreprises et mettant la créativité au cœur de ses valeurs et de ses formations.

CERIS (Centre d'Enseignement et de Recherche en Informatique et Systèmes) est l'un des 3 centres d'enseignements et de recherches de l'IMT Mines Alès qui est spécialisé dans les domaines de l'informatique, de l'intelligence artificielle et de l'industrie du futur. Il réunit aujourd'hui les équipes de recherches I3A (Informatique, Image, Intelligence Artificielle) et ISOAR (Ingénierie des Systèmes des Organisation pour les Activités à Risques), et regroupe les deux départements d'enseignement 2IA (Informatique et Intelligence Artificielle) et PRISM (Performance Industrielle et Systèmes Mécatroniques).

Le projet HUT est donc étudié au sein du CERIS mais aussi de 12 autres laboratoires partenaires tels que DMeM, LICeM, EPSYLON, EuroMOV, LIRMM et Espaces Dev, IES, MRM de l'Université de Montpellier, LIFAM de l'ENSAM (Ecole d'Architecture de Montpellier), LERASS, PRAXILING de l'Université Paul Valéry. De plus, 7 entreprises partenaires contribuent à ce projet telles que DELILED, SYNOX, ENEDIS, IKEA, NEXITY et OceaSoft, et enfin différents partenaires institutionnels participent et aident au financement du projet, tels que l'Université de Montpellier, l'Université Paul Valéry de Montpellier, Montpellier

Méditerranée Métropole, CNRS, MSHSud, IMT Mines Alès, ENSAM, Fondation Université de Montpellier, FEDER (Fond Européen de Développement Régional) et la Région Occitanie. Ces entités ont des rôles particuliers au sein du projet développés en annexes. (cf. *Annexes page 36*).



Figure 3 : Partenaires du projet HUT

## 1.4 Objectifs du stage

Les objectifs de ce stage se distinguaient sous deux catégories.

Premièrement, il s'agissait d'étudier des approches d'intelligence artificielle permettant d'apprendre les habitudes des habitants d'un appartement connecté afin de pouvoir automatiser ou anticiper le déclenchement d'actions ou de scénarios. C'est-à-dire, cela revient à réaliser une étude des données de sorte qu'on puisse réaliser une interprétation contextuelle de celles-ci et établir des séquences d'actions. Celles-ci alimenteraient une intelligence artificielle qui pourrait alors prédire des enchaînements d'actions pour enclencher des appareils.

Secondement, il était souhaité d'étudier le concept de « jumeau numérique ». En effet, l'appartement devant être démantelé à la fin du projet, le développement de jumeaux numérique de ses capteurs pourrait permettre de prolonger au-delà de la fin du projet, les expérimentations réalisées à l'aide de leurs données après la fin du projet initial. Ceci contribuerait aussi au premier objectif.

## 1.5 Démarche établie

Pour répondre à ces objectifs, nous avons choisi dans un premier temps avec mon collaborateur d'explorer le sujet et le projet à travers les réalisations précédentes et une recherche documentaire succincte pour nous ancrer dans celui-ci par rapport à ses problématiques. Ensuite, dans un second temps, nous avons cherché à étudier les données qui nous ont été mises à disposition de manière approfondie, ceci afin de nous les approprier et de réaliser au mieux des interprétations de celles-ci. Enfin, dans un dernier temps, nous avons cherché à réaliser des prédictions en établissant des modèles à partir des données.

Ce rapport présentera alors nos différentes réalisations présentées ci-dessus dans l'ordre indiqué. Je finirai par conclure sur l'avenir de ce projet en partageant mon retour d'expérience sur ce stage.

## 2 Réalisations de ce stage

### 2.1 Introduction

Outre le travail préliminaire d'appropriation du sujet par une lecture et une recherche de travaux précédents réalisés sur le projet ou en lien avec la thématique, les réalisations durant ce stage se distinguent en deux catégories : une analyse approfondie des données et des prédictions de séquences. Vous trouverez ci-dessous ces deux études avec leurs démarches et résultats explicités.

Ensuite, certains choix ont dû être réalisés à travers le projet notamment celui se portant sur le jeu de données à l'étude. En effet, un problème s'est posé dès le départ avec le jeu de données collectés à travers l'appartement observatoire du projet HUT. Ses données sont difficilement exploitables dû à une erreur d'horodatage des données et des fréquences de remontées des données différentes selon les capteurs. L'objectif d'établissement de scénarios se basant sur une chronologie précise, ce jeu de données ne peut pas répondre à cette exigence chronologique. Ainsi, l'étude des données s'est portée sur le jeu de données établi par Gibson Chimamiwa, Marjan Alirezaie, Federico Pecora et Amy Loutfi [1] provenant de l'université d'Orebro en Suède et disponible sur le site de Mendeley Data.

Enfin, par défaut de temps, la thématique des jumeaux numériques n'a pu qu'être partiellement exploré durant ce stage, mais certains éléments réalisés pourront contribuer à son développement prochain. Ce point sera détaillé dans la suite du projet en conclusion.

### 2.2 Résumé du travail préliminaire

Dans un premier temps, nous avons étudié les différents rapports des travaux qui ont pu être réalisés par différents stagiaires nous précédents, notamment le rapport d'Olivier DESSE [3] et de la mission de terrain réalisé au printemps 2022 par Léna DEROCHÉ, Antoine OLISLAGERS et Anna SABOURDY [2]. Ceci a pu nous orienter par rapport aux études qui ont pu être réalisées et les nouvelles analyses que nous pourrions apporter. Comme nous l'avons constaté, ces groupes ont réalisé leur étude sur le deuxième jeu de données issu de Mendeley Data et non du projet HUT. Aussi, ces derniers s'étaient penchés sur les mêmes objectifs qui faisaient l'objet de notre stage. Nous avons remarqué que l'analyse des données avait été moins approfondie que la partie prédiction, c'est pourquoi nous nous sommes attardés un peu plus sur cette première partie.

Nous avons aussi lu et étudié les différents articles de référence du domaine des *smart home*, notamment celui de la thèse de Nadia OUKRICH [4]. Celle-ci se porte plutôt sur la reconnaissance d'activité que la prédiction d'événements. Elle insiste déjà sur la difficulté de cette première tâche qui est le point de départ de la seconde pour mener à des résultats probants. Dans l'analyse des données, nous avons notamment choisi d'étudier certaines combinaisons de capteurs sous forme d'activités, ce qui sera développé par la suite.

Ensuite, nous avons réalisé une recherche succincte à l'aide de mots clés tels que « smart home », « smart home recognition », et leurs équivalents français (« maison connectée », « reconnaissance d'activité »), ceci afin de mieux appréhender les notions du

sujet. Cela nous a permis de distinguer deux notions importantes concernant le sujet qui sont la reconnaissance d'activité et la prédiction d'événements.

La reconnaissance d'activité consiste alors à partir d'un jeu de donnée de reconnaître les actions et les objectifs réalisés par un ou plusieurs utilisateurs, ceci à partir d'une série d'observations unique ou combinées sur l'environnement. Cela requiert des données précises, notamment dans leur chronologie, et surtout une bonne analyse de celles-ci dans un contexte donné, temporel ou spatial.

La prédiction d'événements consiste à prédire les actions d'un ou de plusieurs utilisateurs dans un environnement à partir d'un jeu de données qui retrace les actions précédentes. Une fois de plus, la caractéristique d'un jeu de données précis, notamment dans leur chronologie, est importante pour pouvoir observer et donc prédire l'enchaînement d'action.

Enfin, en lien avec un des objectifs initiaux, nous avons exploré la plateforme OpenHAB qui est utilisée dans la réalisation de jumeaux numériques. Malheureusement, au cours de ce stage, nous n'avons pas été amené à l'utiliser mais certains travaux réalisés pourront être un support pour cet outil. Ceci sera détaillé dans la suite du projet en *Conclusion*.

## 2.3 Analyse des données

### 2.3.1 Présentation du jeu de données

Comme indiqué précédemment, ce n'est pas le jeu de données collecté par HUT qui a été étudié durant ce stage mais un jeu de donnée établi par un autre groupe de recherche. Celui-ci a été collecté pour un appartement où vivait une unique personne dans un environnement connecté selon le protocole comme indiqué dans le rapport associé à ce jeu [4]. Les capteurs utilisés sont les suivants : infrarouge passif, résistances de détections de force, interrupteurs reed, mini-capteurs de lumière à cellules photoélectrique, température, humidité et prises intelligentes. Ainsi, le jeu recense les interactions de l'utilisateur avec son environnement telles que les mouvements au sein des pièces, les pressions appliquées sur le lit et le canapé, l'utilisation d'appareil électroménagers tels que la machine à café, la machine à laver, le lave-vaisselle, le four, la cuisinière, le sandwichmaker, le frigo, la télé, le micro-ondes, etc. Un plan de l'appartement et de la disposition des appareils est disponible en *Annexes (page 37)*.

Il y a trois fichiers CSV (*sensor*, *sensor\_sample\_int*, *sensor\_sample\_float*). Le premier fichier *sensor* contient toutes les données avec comme informations le type de mesure du capteur, le nom de celui-ci avec les id associés. Le deuxième fichier *sensor\_sample\_int* contient les données de types int (entiers), tandis que le troisième fichier *sensor\_sample\_float* contient les données de types float (flottant, nombre à virgule).

Ce jeu de données étant très conséquent en termes de volume, il a fallu diviser celui-ci en différents fichiers ne contenant plus qu'un million de lignes chacun pour ainsi pouvoir ouvrir et exploiter les données. *In fine*, on obtient 187 fichiers pour les données de type int, et 61 fichiers pour les données de type float.

Ensuite, nous avons regroupé dans un même fichier les données concernant un même capteur, ceci afin de réaliser une étude des données par capteurs. Un script déjà réalisé par le doctorant sur le projet **Quentin PEREZ** nous a été fourni pour ceci. En paramètres, nous

choisissions alors le capteur à étudier ainsi que sa fréquence d'échantillonnage sous l'appellation *SAMPLING\_RATE*. Pour les analyses effectuées ici, les données ont été étudiées sous différentes fréquences d'échantillonnage : toutes les minutes et toutes les secondes.

### 2.3.2 Objectifs et démarche

Nous avons choisi de porter une attention importante à l'analyse des données, en faisant une analyse approfondie « à la main » de manière à déceler des généralités. Ceci répond à l'objectif global qui est celui de pouvoir repérer des activités pour ensuite les détecter. On s'intéresse donc à comprendre le comportement individuel et combiné des capteurs pour déceler des généralités, étudier des activités et/ou en reconnaître. Ce genre de jeu de données n'ayant pas été largement exploité dans la communauté de recherche dans ce domaine, cela permet aussi de se rendre compte des limites et des éventuels problèmes avec celui-ci

Pour cela, nous avons mené différentes analyses que nous nous sommes réparties avec mon collègue. Pour ma part, j'ai mené une première analyse visuelle, ceci afin d'observer dans un premier temps le comportement des données capteur par capteur pour déterminer les seuils et les interprétations contextuelles associées, les associer aux pièces, référencer les périodes de collecte et éventuellement remarquer un pattern pour un capteur en particulier. Dans un second temps, j'ai effectué différentes analyses par « activités » pour lesquelles il pouvait y avoir un unique ou plusieurs capteurs combinés à l'étude, ceci pour déterminer les plages d'utilisation et les moments d'utilisation, dans la journée ou la semaine. Ces deux parties ont été effectuées sur l'ensemble des données échantillonnées toutes les minutes.

Quant à mon collègue, il a effectué des analyses pour des données sur une plus courte période (3 semaines) et celles-ci échantillonnées à la seconde. Il a étudié des activités intitulées « permanentes » (c'est-à-dire les capteurs qui ont un comportement continu) et des activités intitulées « instantanées » (données au comportement analogique). Les résultats de ses analyses seront présentés dans son propre rapport.

Ces deux démarches, particulières pour deux échantillons de données différents, permettent alors de mener une analyse approfondie et de reconnaître si des généralités peuvent être portées sur celles-ci, ou bien si on décèle les mêmes biais.

### 2.3.3 Analyse graphique des données

Comme explicité ci-dessus, j'ai réalisé des analyses graphiques. Bien sûr, les conclusions et remarques obtenues ont été partagées avec mon collègue pour la progression du projet.

Dans un premier temps, j'ai réuni les informations globales et importantes pour la totalité des capteurs. (cf. *Annexes page 38 à 40*), où j'ai associé chaque capteur à la pièce correspondante, en référençant les plages de valeurs pour lesquelles les données ont été collectées, ainsi que les interprétations de l'état des capteurs grâce à la détermination de seuils par l'analyse visuelle.

A partir du tableau présent ci-dessous, nous pouvons remarquer certaines ambiguïtés sur les données :

- Existence de valeurs intermédiaires pour des appareils qui ont des états binaires *a priori*. Exemple : capteur de lumière avec lumière éteinte / lumière allumée. Il n'y a pas d'indication supplémentaire dans le rapport du jeu de données pour expliciter ceci.
- Existence de valeurs aberrantes qui peuvent être, par exemple, 10 fois supérieures aux autres maximum « habituels ». Exemple de la machine à laver.
- Modification de seuils selon la plage de données. Exemple du canapé ou du lit.
- Difficulté d'interprétation des données sur l'unité de la donnée renvoyée, sans explicitation dans le rapport du jeu de donnée. Exemple de la balance de pesée.
- Apparente défaillance du capteur ou protocole annoncé de l'occupation de l'appartement par un unique habitant non respecté. Exemple du capteur de mouvement de la salle de bain à partir du 2020-03-18.

C'est bien cette dernière remarque qui nous a alors confirmé le besoin de mener une analyse sur une durée plus courte, telle que 3 semaines, qui a été faite par mon collègue.

Ensuite, dans un second temps, j'ai mené une analyse purement graphique sur la présence dans l'appartement avec des événements spécifiques tels que « concordance de l'entrée dans l'appartement avec la présence dans le couloir (et les sorties) » ; « présence globale » ; « succession d'entrées / présence dans les pièces » et « concordance d'activités ».

Un tableau plus complet est présent en *Annexes (page 40 à 45)* pour étayer les observations, ainsi que des captures graphiques (*images page 45 à 47*). Certaines observations ont été faites avec un échantillonnage plus précis tel que 15 secondes pour venir confirmer ou infirmer des observations faites sur un échantillonnage de 1 minute. Ci-dessous, vous trouverez les conclusions tirées pour chaque événement.

#### **Concordance de l'entrée dans l'appartement avec la présence dans le couloir (et les sorties)**

La corrélation de ses deux capteurs peut indiquer les événements suivants :

- L'habitant entre dans l'appartement : ouverture de la porte suivie de mouvement dans l'entrée
- L'habitant sort de l'appartement : ouverture de la porte suivie d'absence de mouvement dans l'entrée

D'autres événements sont à prendre en compte :

- Quelqu'un est invité dans l'appartement et / ou l'habitant vérifie la fermeture de sa porte d'entrée (en ouvrant et fermant la porte) : ouverture et fermeture de la porte très rapprochées et présence (présence importante si éventuellement plus d'une personne) dans l'entrée
- L'habitant ouvre la porte et discute sur le perron sans inviter la personne à rentrer : ouverture et fermeture de la porte rapprochées, mouvements **possibles mais non nécessaires** entre ces ouvertures / fermetures

Critique de l'analyse :

- Les données étant extraites à la minute, certaines valeurs (pics) établissant la corrélation des résultats graphiquement pourraient être absentes et fausser l'analyse à un instant t.
- En restant sur l'hypothèse principale de la présence d'une unique personne dans l'appartement, ce sont bien les capteurs de mouvement des différentes pièces qui indiquent la présence ou non de l'habitant dans l'appartement (hypothèse vérifiée au point n°2 « Présence globale »).

L'analyse de la présence globale pourra nous indiquer si cette analyse de mouvement dans l'entrée et de la porte est suffisante pour tirer les conclusions exposées ci-dessus.

### Présence globale

Il semble difficile de tirer des conclusions malgré le fait de tirer la précision des analyses jusqu'à une remontée des informations à 15 secondes. Voici ce qui est constaté :

- Absence de mouvement la nuit ou quelques mouvements, principalement dans la chambre et dans la salle de bain. On peut en convenir que la personne bouge la nuit quand elle dort ou bien qu'elle se lève pour aller dans la salle de bain. Aucune présence n'est détectée, la plupart du temps, dans ce cas dans le couloir alors que c'est un passage obligé entre chaque pièce.
- Relevé d'incohérences sur certaines activités compte-tenu de la temporalité : par exemple, grande activité dans la salle de bain (des heures), pendant la nuit, sans détection dans le couloir pour le changement de pièce par la suite. On peut suspecter qu'un problème est survenu sur le capteur.
- L'habitant ne semble pas sortir de l'appartement : les détections de la porte d'entrées sont quasi-toujours suivi de mouvements dans d'autres pièces plusieurs minutes après l'action sur la porte, sans détection dans le couloir (incohérence des résultats ou manque de données). On relève aussi des contacts avec la porte avec détection de l'habitant dans différentes pièces avant et après cette action. Trois hypothèses envisagées :
  - o L'habitant reçoit quelqu'un sur le perron de sa porte mais ne sort pas
  - o L'habitant reçoit quelqu'un chez lui et donc ne sort pas, mais c'est bien l'invité qui entre et qui sort
  - o Il y a plusieurs personnes qui vivent dans l'appartement (contrairement à ce que le rapport sur le dataset indique).

Il y est difficile de déterminer quelle est la situation ici.

Je suis allée jusqu'à échantillonner les données toutes les 15 secondes pour l'analyse de la présence (et surtout des entrées/sorties), afin d'être le plus précis possible et de manquer le moins de données possibles. Il est possible que la certaine incohérence relevée, notamment des ouvertures/fermetures de portes peuvent être manquées, ce qui pourrait expliquer certaines situations. Il n'est pas possible de traiter les données avec une fréquence d'échantillonnage plus faible, sinon de le faire sur une quantité de données restreinte.

Les données étant remontées à la seconde pour chaque capteur, il peut s'agir aussi d'un problème dans la duplication de la donnée à la minute  $m+1$ , en particulier pour les capteurs de détection de présence, qui ne correspond pas en réalité à l'état à cet instant. On peut, en dernier lieu, supposer aussi une défaillance technique des capteurs qui n'auraient pas relevé l'information d'une présence à un instant  $t$  particulier durant lequel il y aurait eu une présence, et vis-versa.

### Succession d'entrées / présence dans les pièces

Il est difficile de tirer des conclusions par l'analyse graphique. Une analyse plus quantitative est donc plus pertinente pour tenter de tirer des généralités. Les *a priori* qu'on pourrait avoir sur un comportement « normal » ou « stéréotype » sont à manier avec précaution.

### Concordance d'activités

Cette partie correspondait à des vérifications sur certains appareils pour déceler des présences qui étaient *a priori* jugées anormales. On en conclut une certaine défaillance, problème dans la duplication des données ou une fragilité pour les capteurs de mouvement, ce qui avait déjà été observé un peu plus tôt.

Tableau 1 : Conclusions des observations graphiques menées

Les analyses graphiques ayant montrées leurs limites, j'ai donc choisi de poursuivre en menant une analyse plus « quantitative » afin de déceler des généralités sur les données. On a pu se rendre compte avec cette première approche de la difficulté du traitement des données relevées par les capteurs de mouvement pour lesquelles on trouve ici beaucoup d'incohérences.

## 2.3.4 Analyse des données sous forme d'activités

Pour une analyse plus « quantitative », j'ai choisi d'étudier les capteurs sous forme d'activités, celles-ci pouvant correspondre à un capteur unique ou une combinaison de capteurs. Il s'agit d'activités dans un premier temps basiques telles que : « regarder la télé » ; « dormir dans le lit » ; « faire un repas » ; et « activité de la salle de bain ».

Pour ces quatre parties, un même schéma d'analyse a été réalisé. Dans un premier temps, j'ai relevé les données du capteur pour lequel celui-ci est considéré en « état de marche ». Ensuite, j'ai réalisé les plages d'utilisation entre ces différents instants relevés. J'ai corrigé ces plages pour lesquelles la différence entre la fin et le début d'une autre plage est inférieure à un écart fixé, ce qui nous donne les plages finales utilisées à l'étude.

J'ai étudié alors celles-ci en observant pour l'ensemble des données sur une journée, les fréquences et les durées moyennes d'utilisation selon un créneau horaire particulier. J'ai aussi réalisé cette étude sur la semaine.

Les résultats seront présentés ici en quatre parties distinctes correspondant à chacune des activités citées ci-dessus.

### 2.3.4.1 Activité « regarder la télé »

Utilisation des capteurs suivants :

- Capteur de pression du canapé (réf : P01 - 5889)
- Capteur de lumière de la télé (réf : L01 - 6127)

*Problèmes relevés dans cette partie :*

La plage du 2020-02-26 au 2020-03-17 n'est pas suffisante pour l'étude de ce capteur. Il faut donc allonger la plage jusqu'au 2020-06-01 (passé cette date, on décèle un comportement anormal du capteur de pression du canapé).

*Choix :*

- Dans un premier temps, il a été envisagé de travailler aussi avec le capteur de mouvement de la pièce concernée. Au cours de l'étude, il a été décidé de ne pas étudier cette combinaison de capteurs ici. En effet, selon le fonctionnement des capteurs de mouvement explicité dans le rapport du jeu de données associé [4], ces derniers ne marquent une présence que si un mouvement est détecté. Or, lors de la phase de sommeil, le sujet peut être assis sur le canapé, ne pas bouger mais être bien présent dans la pièce, sans que le capteur ne le détecte.

*Méthodologie :*

Ici, j'ai basé cette activité sur les données provenant de deux capteurs (pression du canapé et lumière de la télé). La méthodologie suivie est celle qui a été explicitée en introduction à cette partie. L'écart fixé pour la réunion des plages est de 10 minutes. A la réunion de ces plages, nous en comptons 75 pour cette activité.



Les différentes figures et tableaux présentés ci-dessous sont disponibles dans une version plus agrandie en *Annexes (page 48 à 50)*.

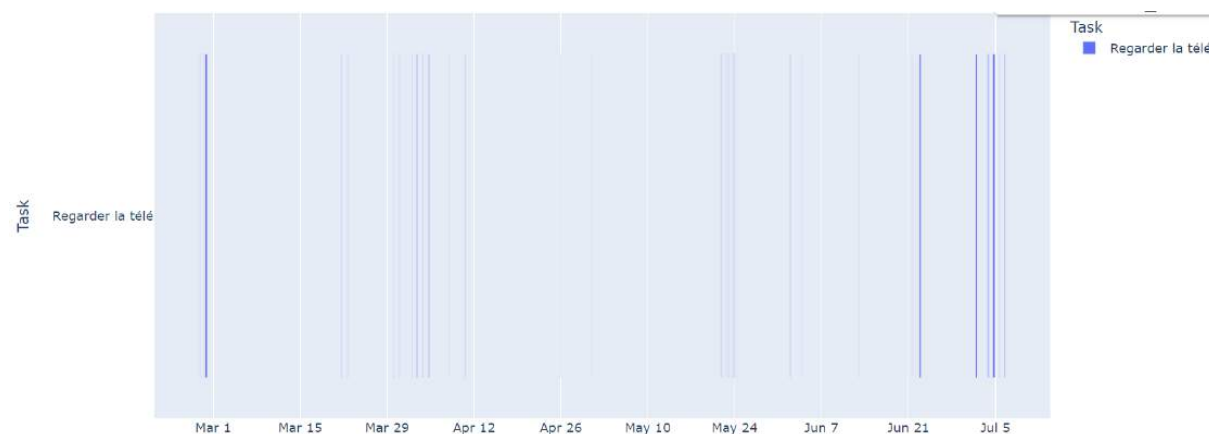


Figure 4 : Visualisation des plages de l'activité "regarder la télé" pour l'ensemble des données

### Résultats :

Parmi les analyses que j'ai réalisées, je me suis intéressée à la fréquence de l'activité de différentes manières.

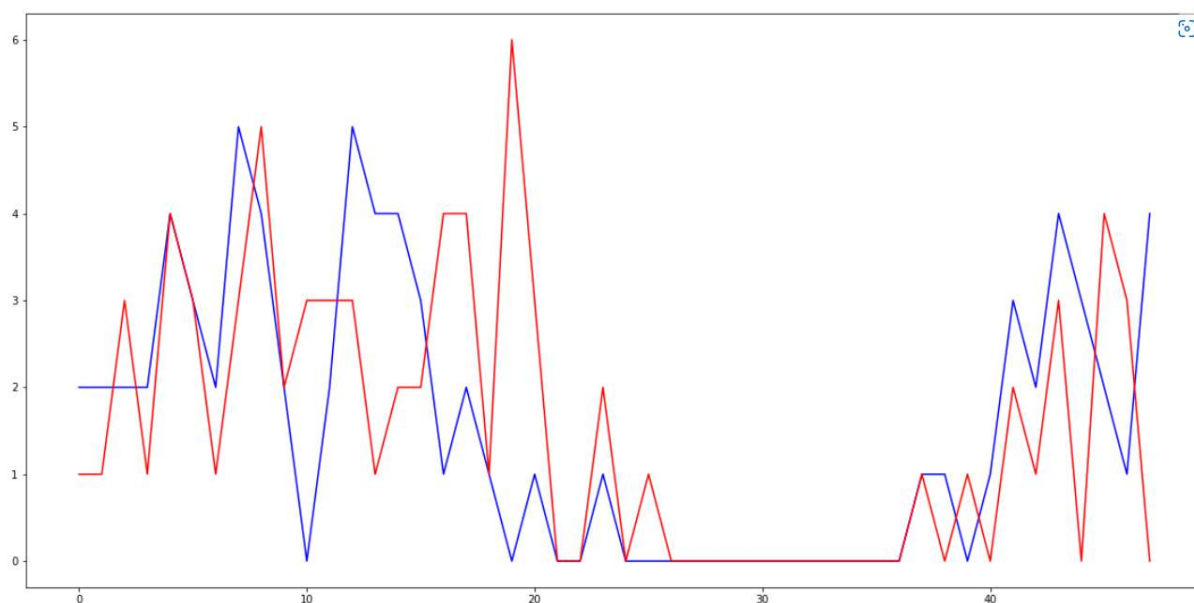


Figure 5 : Visualisation de la fréquence selon le début et la fin de l'activité "regarder la télé"

N.B. : L'indice 0 du graphe correspondant à 15:00, le graphique couvrant 24 heures.

Sur le graphique, on observe en bleu les fréquences pour le début de l'activité et en rouge pour la fin. L'axe des abscisses est numéroté de 0 à 47 correspondant à l'échelle de 15:00 à 14:30, représentant tous les créneaux de la journée toutes les 30 minutes, chaque créneau étant attribué pour une donnée dont l'heure correspond à ce créneau à 15 minutes d'intervalle de part et d'autres de l'horaire du créneau.

Ce qu'on remarque surtout, ce sont des périodes de non-activité, c'est-à-dire que l'activité n'est jamais lancée ou finie sur cette période :

- Fréquence nulle de l'indice 24 à 36 soit de 03:00:00-09:00:00, ce qui ne paraît pas incohérent puisqu'il s'agit de la nuit.
- Cette plage peut s'allonger avec les indices 22 à 34 dont la fréquence est nulle ou faible (unitaire), c'est-à-dire à partir de 02:00 du matin, ce qui n'est pas incohérent ; mais aussi avec la plage des indices 35 à 39 soit 8:30:00-10:30:00 où idem la fréquence est nulle ou faible (unitaire).

Cela ne confirme pas si la télé était allumée ou non, mais on observe parmi les fréquences de début ou fin de l'activité "regarder la télé" que celle-ci n'est pas exercée dans la matinée, mais dans l'après-midi, et plutôt dans la soirée.

Ensuite, on réalise les études suivantes :

- Etude de la durée des activités
- Etude du moment de la journée de l'activité (sur la durée) par rapport au jour de la semaine.

Pour cela on va s'intéresser aux plages d'activité qui durent plus de 20 minutes et on obtient 50 plages. On va calculer la moyenne de la durée de l'activité sur un créneau, les créneaux étant répertoriés de la même manière que présenté précédemment, uniquement que pour les dates et horaires de début des plages. Nous récupérons aussi le nombre de données répertoriées sur ce créneau c'est-à-dire la fréquence de début de l'activité. On obtient alors ce type de table :

créneau	durée_moy	fréquence		créneau	durée_moy	fréquence	
0	0:00:00	NaN	0	17	8:30:00	NaN	0
1	0:30:00	NaN	0	18	9:00:00	NaN	0
2	1:00:00	1:24:00	1	19	9:30:00	NaN	0
3	1:30:00	NaN	0	20	10:00:00	0:24:00	1
4	2:00:00	NaN	0	21	10:30:00	NaN	0
5	2:30:00	0:47:00	1	22	11:00:00	0:34:00	1
6	3:00:00	NaN	0	23	11:30:00	2:39:30	2
7	3:30:00	NaN	0	24	12:00:00	0:22:00	1
8	4:00:00	NaN	0	25	12:30:00	1:34:30	2
9	4:30:00	NaN	0	26	13:00:00	3:28:00	1
10	5:00:00	NaN	0	27	13:30:00	NaN	0
11	5:30:00	NaN	0	28	14:00:00	4:35:00	2
12	6:00:00	NaN	0	29	14:30:00	0:58:40	3
13	6:30:00	NaN	0	30	15:00:00	3:23:00	1
14	7:00:00	NaN	0	31	15:30:00	4:28:01	1
15	7:30:00	NaN	0	32	16:00:00	5:11:00	2
16	8:00:00	NaN	0	33	16:30:00	0:34:30	2
				34	17:00:00	2:33:00	3
				35	17:30:00	4:00:00	2
				36	18:00:00	2:42:00	1
				37	18:30:00	1:48:00	4
				38	19:00:00	1:42:00	4
				39	19:30:00	0:44:00	2
				40	20:00:00	NaN	0
				41	20:30:00	NaN	0
				42	21:00:00	2:45:40	3
				43	21:30:00	1:03:20	3
				44	22:00:00	1:56:00	4
				45	22:30:00	1:47:00	3
				46	23:00:00	NaN	0
				47	23:30:00	NaN	0

Figure 6 : Table des fréquences et durées moyennes pour l'activité "regarder la télé"

On observe globalement une grande inactivité de l'activité "regarder la télé" pendant la période de la nuit (de 23:00:00 à 10:30:00), c'est-à-dire que l'activité n'est pas lancée sur ces créneaux, hormis quelques exceptions qui ne semblent être que des événements ponctuels (arrivés une seule fois).

L'activité se trouve alors plutôt dans la deuxième partie de la journée que l'on peut catégoriser en 3 sous-parties : midi, après-midi, soirée. Une analyse plus précise selon les jours de la semaine pourrait apporter plus d'informations, ce qui donne le type de table suivante :

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim
0:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
0:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
1:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	1:24:00 - 1
1:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
2:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
2:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:47:00 - 1	Nan - 0
3:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
3:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
4:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
4:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
5:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
5:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
6:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
6:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
7:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
7:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
8:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
8:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
9:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
9:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
10:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:24:00 - 1
10:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
11:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:34:00 - 1	Nan - 0	Nan - 0
11:30:00	3:29:01 - 1	Nan - 0	Nan - 0	Nan - 0	1:50:00 - 1	Nan - 0	Nan - 0
12:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:22:00 - 1	Nan - 0	Nan - 0
12:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	1:28:01 - 1	1:41:00 - 1
13:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:28:00 - 1	Nan - 0	Nan - 0
13:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
14:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	8:39:01 - 1	0:31:00 - 1
14:30:00	Nan - 0	Nan - 0	Nan - 0	1:40:00 - 1	Nan - 0	Nan - 0	0:38:00 - 2
15:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:23:00 - 1	Nan - 0	Nan - 0
15:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	4:28:01 - 1	Nan - 0
16:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	5:11:00 - 2	Nan - 0	Nan - 0
16:30:00	0:47:00 - 1	Nan - 0	Nan - 0	0:22:00 - 1	Nan - 0	Nan - 0	Nan - 0
17:00:00	Nan - 0	Nan - 0	1:53:00 - 1	2:53:00 - 2	Nan - 0	Nan - 0	Nan - 0
17:30:00	7:22:00 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:38:00 - 1
18:00:00	Nan - 0	Nan - 0	Nan - 0	2:42:00 - 1	Nan - 0	Nan - 0	Nan - 0
18:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	4:28:00 - 1	1:09:00 - 2	0:26:00 - 1
19:00:00	Nan - 0	Nan - 0	5:25:00 - 1	0:25:00 - 1	0:30:00 - 1	Nan - 0	0:28:00 - 1
19:30:00	Nan - 0	Nan - 0	Nan - 0	0:54:00 - 1	0:34:01 - 1	Nan - 0	Nan - 0
20:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
20:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
21:00:00	3:57:01 - 1	Nan - 0	Nan - 0	Nan - 0	0:40:00 - 1	3:40:00 - 1	Nan - 0
21:30:00	Nan - 0	Nan - 0	0:35:00 - 1	1:45:00 - 1	Nan - 0	Nan - 0	0:50:00 - 1
22:00:00	2:40:00 - 1	Nan - 0	2:27:00 - 1	Nan - 0	Nan - 0	0:56:00 - 1	1:41:01 - 1
22:30:00	Nan - 0	2:39:00 - 1	Nan - 0	0:59:00 - 1	1:43:00 - 1	Nan - 0	Nan - 0
23:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
23:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0

A noter pour rappel que ces résultats sont sur la base des données de 2020-02-26 à 2020-07-06.

Une chose que l'on trouve commune aux jours est l'importance de l'activité dans la période de la soirée. On remarque aussi que les quelques données enregistrées pour un horaire plutôt nocturne se passent le week-end (hormis une donnée).

On note aussi que lorsque l'activité se déroule l'après-midi, il s'agit plutôt du week-end, mais aussi du vendredi. Il y a tout de même des activités relevées des jours de semaine, mais à noter qu'il s'agit la plupart du temps d'une occurrence unique sur la plage de données étudiée.

De manière plus globale, on remarque que 45 plages sont distribuées dans le calendrier, ceci à partir de 50 plages distinctes. 90% des plages du calendrier sont donc constituées à partir de plages uniques et non d'une moyenne sur plusieurs plages. Dans ces conditions, il peut être considéré présomptueux de tirer des conclusions sur des habitudes précises. Ceci est dû aux simplifications réaliser plus haut, notamment la non-considération des plages inférieures à 20 minutes. Cette simplification a retiré une vingtaine de plages. Il peut être intéressant d'observer l'effet de cette non-considération de données en ajoutant ces plages à l'étude, mais ce choix a été motivé du fait qu'il y avait des plages de très faible durée qui pouvait alors fausser le calcul de la moyenne.

La conclusion la plus probante qui peut être retirée ici est l'activité importante dans la soirée de l'activité "regarder la télé". Ceci appuierait l'hypothèse faite pour un scénario (activité "regarder la télé" qui précéderait l'activité "dormir dans le lit").

Figure 7 : Calendrier des fréquences et durées moyennes pour l'activité "regarder la télé"

### 2.3.4.2 Activité « dormir dans le lit »

Utilisation des capteurs suivants :

- Capteur pression du lit (P02 - 5896)
- Capteur de mouvement sous le lit (M04 - 6686)
- Capteur de mouvement dans la pièce (bedroom) (M03 - 5892)

Les capteurs de mouvement du dessous du lit et de la pièce ne semblent pas pour assez fiables ou consistantes pour porter des analyses concrètes comme a pu le montrer la visualisation précédente, notamment pour les raisons ci-dessous :

- Possibilité d'étudier la combinaison des données qu'à partir du 2020-03-24 dû au fait que le capteur de mouvement sous le lit ne contient pas de données avant cette date.
- Problèmes du capteurs relevés pour les capteurs de mouvement à partir du 2020-04-20 au 2020-06-06 pour celui de la pièce (bedroom) et du 2020-03-24 au 2020-03-26 pour le capteur de mouvement sous le lit (moins gênant)
- Absence de détection de mouvement sous le lit avec présence sur le lit dû au fonctionnement même du capteur (même justification que pour l'activité précédente)

On remarque que les paliers distinguant l'état de l'activité c'est-à-dire présence détectée sur le lit et absence de présence détectée sur le lit, diffèrent selon la plage de dates étudiée.

*Choix :*

- L'activité "dormir dans le lit" sera étudiée avec un unique capteur : le capteur de pression du lit
- Réajustement manuel du seuil indiquant l'état de l'activité selon la date étudiée

*Méthodologie :*

On va appliquer la même démarche que pour l'activité précédente. Ainsi, on obtient 39 381 plages que j'ai réduit à 471 plages en les réunissant. (*Figure 19* disponible en *Annexes page 50*).

*Résultats :*

De nouveau, parmi les analyses que j'ai réalisées, je me suis intéressée à la fréquence de l'activité de différentes manières. La figure (*Figure 20 page 50*) est disponible en *Annexes* et les remarques sur celles-ci sont explicitées ci-dessous.

La figure correspond à la plage de dates suivante : du 2020-02-26 à la fin des données (2020-08).

En bleu on observe la fréquence du coucher et en rouge la fréquence du lever. On observe que l'étendue sous la courbe de cette dernière est plus élargie que pour le coucher. Les fréquences les plus importantes relevées sont les suivantes :

- Un pic distinctif du lever à l'indice 41 ce qui correspond à 11:30:00, à l'indice 40 ce qui correspond à 11:00:00
- Pour le lever, des pics moins importants sont présents aux indices 38 (10:00:00), 42 (12:00:00), 39 (10:30:00), ainsi que 6 (18:00:00)

- Pour le coucher, on observe que les fréquences sont plus étalées. On observe le plus haut pic à l'indice 15 et 17 ce qui correspond à 22:30:00 et 23:30:00, puis à l'indice 16 ce qui correspond à 23:00:00, et l'indice 14 et 18 ce qui correspond à 22:00:00 et 00:00:00.
- Les pics moins importants mais non négligeables concernant le coucher sont les suivants : à l'indice 13 et 19 (21:30:00 et 00:30:00).

Sur cette durée, 471 plages ont été identifiées avec un nettoyage des données supprimant les plages inférieures à 1 minute et réunissant les plages dont l'écart est inférieur ou égal à 10 minutes. On peut statuer les pourcentages suivants sur les probabilités des activités suivantes :

- L'habitant se lève aux alentours de la plage 11:00:00-11:30:00 à 15 minutes d'intervalles : 13,4%
- L'habitant se lève aux alentours de la plage 10:00:00-12:00:00 à 15 minutes d'intervalles : 27,8%
- L'habitant va se coucher aux alentours de la plage 22:30:00-23:30:00 à 15 minutes d'intervalles : 16,1%
- L'habitant va se coucher aux alentours de la plage 22:30:00-00:30:00 à 15 minutes d'intervalles : 25,3%

Tout comme sur la première étude, on peut fortement supposer qu'il y a d'autres activités en amont du coucher qui sont corrélées ou non à cette activité "dormir sur le lit".

Pour les analyses suivantes, on s'intéresse aux plages inférieures ou égales à 2 heures ce qui réduit les données à 174 plages.

Vous trouverez en *Annexes* le tableau (*Tableau 7 page 51*), qui présente les fréquences et durées moyennes pour le début de cette activité selon des créneaux de 30 minutes réparties sur la journée et ci-dessous l'explicitation de ces résultats.

On observe que l'activité se déroule principalement le soir (de 19:30:00 à 23:30:00) mais on remarque aussi étrangement que cette activité débute en plein milieu de la nuit (début entre 00:30:00 et 04:00:00). On peut expliquer ceci avec un habitant qui peut parfois se coucher très tard, ou des réveils dans le cours de la nuit.

Une analyse semainière pourra donner plus de détails pour connaître les habitudes du coucher qui est illustrée par un tableau (*Tableau 8 page 52*) disponible en *Annexes* et dont les analyses sont les suivantes.

On retrouve une activité importante en fin de soirée avec un coucher plutôt vers 21:30:00-22:00:00 ; mais on observe toujours, répartie dans la semaine un début d'activité important dans le milieu de la nuit.

En notant la fréquence d'utilisation de l'activité, on remarque que l'habitant a plutôt l'habitude de se coucher dans les créneaux horaires de 22:30 à 23:30 ou de 00:30 à 01:30; le 01:30 étant le plus souvent atteint le samedi c'est-à-dire dans la nuit de vendredi à samedi (début du week-end). On remarque aussi un horaire tardif important encore dans la nuit de vendredi à samedi à 03:00 mais aussi dans la nuit de samedi à dimanche à 02:30.

#### 2.3.4.3 Corrélation des activités « regarder la télé » et « dormir dans le lit » :

Suite à l'étude des plages horaires de ces activités, notamment par le graphique de ces plages, on peut envisager qu'il y a une corrélation entre ces activités, c'est-à-dire un scénario qui respecte l'hypothèse suivante : *L'habitant regarde la télé avant d'aller se coucher*

Pour vérifier cette hypothèse, il faut alors regarder selon la période de données disponibles et considérées « fiables » : du 2020-02-26 au 2020-07-06.

	début_rgtv	fin_rgtv	durée_rgtv	début_sleep	fin_sleep	durée_sleep
0	2020-02-27 18:47:00	2020-02-27 19:12:00	0:25:00	2020-02-27 19:15:00	2020-02-27 19:22:01	0:07:01
1	2020-02-27 21:38:00	2020-02-27 22:49:00	1:11:00	2020-02-27 22:53:00	2020-02-28 10:04:57	11:11:57
2	2020-03-29 21:52:00	2020-03-29 23:33:01	1:41:01	2020-03-29 23:36:00	2020-03-29 23:38:00	0:02:00
3	2020-04-01 22:08:00	2020-04-02 00:35:00	2:27:00	2020-04-02 00:44:00	2020-04-02 09:47:00	9:03:00
4	2020-04-02 22:33:00	2020-04-02 23:32:00	0:59:00	2020-04-02 23:37:00	2020-04-03 10:45:00	11:08:00
5	2020-04-03 16:13:00	2020-04-03 19:11:00	2:58:00	2020-04-03 19:12:00	2020-04-03 19:25:01	0:13:01
6	2020-04-03 20:46:00	2020-04-03 21:26:00	0:40:00	2020-04-03 21:27:00	2020-04-04 12:19:00	14:52:00
7	2020-04-04 20:59:00	2020-04-04 21:08:00	0:09:00	2020-04-04 21:23:00	2020-04-05 12:17:01	14:54:01
8	2020-04-07 22:27:00	2020-04-08 01:06:00	2:39:00	2020-04-08 01:19:00	2020-04-08 10:52:58	9:33:58
9	2020-04-26 20:52:00	2020-04-26 20:55:00	0:03:00	2020-04-26 20:59:00	2020-04-27 11:40:00	14:41:00
10	2020-05-21 21:23:00	2020-05-21 23:08:00	1:45:00	2020-05-21 23:24:00	2020-05-21 23:32:00	0:08:00
11	2020-05-21 23:20:00	2020-05-21 23:22:00	0:02:00	2020-05-21 23:24:00	2020-05-21 23:32:00	0:08:00
12	2020-05-22 10:59:00	2020-05-22 11:33:00	0:34:00	2020-05-22 11:37:53	2020-05-22 12:31:00	0:53:07
13	2020-05-23 02:38:00	2020-05-23 03:25:00	0:47:00	2020-05-23 03:30:00	2020-05-23 12:10:00	8:40:00
14	2020-05-24 00:57:00	2020-05-24 02:21:00	1:24:00	2020-05-24 02:25:00	2020-05-24 02:27:59	0:02:59
15	2020-06-01 21:06:59	2020-06-02 01:04:00	3:57:01	2020-06-02 01:07:00	2020-06-02 12:41:00	11:34:00
16	2020-06-03 17:05:00	2020-06-03 18:58:00	1:53:00	2020-06-03 19:03:00	2020-06-03 19:15:00	0:12:00
17	2020-06-22 17:21:00	2020-06-23 00:43:00	7:22:00	2020-06-23 00:49:00	2020-06-23 10:17:01	9:28:01
18	2020-07-01 19:06:00	2020-07-02 00:31:00	5:25:00	2020-07-02 00:41:58	2020-07-02 07:32:39	6:50:41
19	2020-07-04 13:59:00	2020-07-04 22:38:01	8:39:01	2020-07-04 22:42:00	2020-07-04 22:49:00	0:07:00
20	2020-07-05 14:15:00	2020-07-05 14:46:00	0:31:00	2020-07-05 14:48:00	2020-07-05 15:03:00	0:15:00
21	2020-07-05 15:07:00	2020-07-05 16:12:00	1:05:00	2020-07-05 16:19:00	2020-07-05 16:50:00	0:31:00

Tableau 2 : Table des plages des deux activités "regarder la télé" et "dormir dans le lit" qui se suivent à 20 minutes

N.B. : disponible en version agrandie en Annexes page 53

Pour rappel, il s'agit des plages à partir des données de 2020-02-26 à 2020-07-06.

Cette concordance pour le scénario a été établi pour le scénario *L'habitant regarder la télé puis va se coucher* pour un écart considéré de 20 minutes entre les deux activités. On obtient 22 plages.

10 de ces plages, soit la moitié, sont significatives pour ce scénario. 16 de ces plages concernent des activités qui se déroulent dans la soirée (soit un peu plus de 75%). Cela confirme les hypothèses émises telles que la corrélation de ces deux activités se déroulant le soir. Néanmoins, pour une étendue de la plage de données sur 4 mois, nous obtenons peu de plages, cela étant dû au faible nombre de plages de départ concernant l'activité « regarder la télé ». On retrouve ici la moitié de ces plages correspondant à cette activité.

#### 2.3.4.4 Activité « faire un repas »

Utilisation des capteurs suivants :

- Détecteur de contact de la porte du frigo (6253 - D03)
- Courant de la machine à café (6632 - SMP01)
- Courant du sandwichmaker (6633 - SMP03)
- Courant du lave-vaisselle (6634 - SMP04)
- Courant de la bouilloire (6635 - SMP04)
- Courant du micro-onde (6696 - SMP06)
- Lumière de la cuisinière (5887 - L02)
- Détecteur de mouvement de la pièce cuisine (5893 - M05)

On va appliquer la même méthode que pour les études précédentes. Ainsi, on obtient le nombre de plages (après réunion) suivant :

- Frigo : 315 plages
- Café : 76 plages
- Sandwichmaker : 58 plages
- Lave-vaisselle : 159 plages (réduit de nouveau à 150 en enlevant les plages de moins de 1 minute).
- Bouilloire : 5 plages
- Micro-onde : 28 plages
- Lumière de la cuisinière : 467 plages (réduit de nouveau à 393 plages en enlevant les plages de moins de 1 minute).

Le nombre de plages pour la bouilloire étant très faible, celle-ci n'a pas été étudiée. De plus, le graphique affichant les plages sur la durée n'est pas exposé ici car il est peu visible dû à la densité de plages très différente selon les capteurs sur tout le long de la collecte de données, et à la durée de ces plages (parfois juste 1 seconde).

On réalise une première étude en combinant 3 capteurs (le frigo, le sandwichmaker et le micro-onde) qui sont les appareils qui correspondent le mieux à l'activité « faire un repas ».

On analyse dans un premier temps la fréquence de cette activité (en considérant simplement le début de l'activité, comme réalisé précédemment). L'illustration se trouve en *Annexes (Figure 21 page 54)* et l'explicitation de ce dernier ci-dessous.

Pour rappel, ici nous étudions uniquement trois appareils pour l'activité : "faire un repas". On remarque les choses suivantes :

[1. Pourcentages par rapport aux nombres de plages : 401]

{2. Pourcentages par rapport aux nombres de créneaux totaux : 48}

- On retrouve certains créneaux communs pour lesquels les fréquences d'utilisation sont élevées : à 10:30 (29) [soit 7,2%], 11:30 (27) [soit 6,7%], 18:30 (22) [soit 5,5%], 11:00 (19) [soit 4,7%], 14:00 (16) [soit 4%], 12:30 / 20:00 (15) [soit 7,5%], 17:30 (14) [soit 3,5%], ce qui fait un total de 39%.
- Les pics en communs les plus élevés correspondent à la fin de matinée : 10:30, 11:30, 11:00.
- On observe de nouvelles valeurs avec des fréquences élevées, dû à l'utilisation plus importante du frigo sur ces périodes : 10:30 / 18:00 (17 pour chaque) [soit 8,5%], 9:00 (16) [soit 4%] et 15:30 / 17:00 (14 pour chaque) [soit 7%], ce qui fait un total de 58,6% de l'activité avec les plages précédentes. Ces nouvelles valeurs ne sont pas contradictoires avec les conclusions apportées lors de la première analyse.
- On remarque une absence d'utilisation ou une faible fréquence d'utilisation toujours sur les plages de 03:00 à 6:30 {soit 16,7% de créneaux vides}, puis de manière étendue de 01:30 à 07:30 {soit 27% de créneaux vides}. Pour une deuxième analyse (disponible en *Annexes*), où les plages de 1 seconde du frigo n'étaient pas considérées, on retrouve ici une base commune et notamment le fait que l'activité « faire un repas » n'est pas réalisée la nuit.

A noter que pour cette analyse, les données du frigo sont très importantes ici (volume de 300 données). Une étude n'incluant pas les données du frigo dans cette activité a aussi été menée et est disponible en *Annexes page 54*.

On réalise une deuxième étude, cette fois-ci simplement sur l'appareil de la machine à café. Trois figures illustrent les explications ci-dessous (*Figure 22, Tableau 10 et Tableau 11 page 55 et 56*).

On observe une cohérence dans l'allure du graphe avec une fréquence d'utilisation plutôt aux moments de la fin de matin et midi, ce qui correspond aux stéréotypes d'utilisation. On note surtout l'absence d'utilisation de la machine à café de 20:00:00 à 6:30:00 (la nuit), avec des pics d'utilisation en milieu de journée (sur le midi).

On réalise une troisième étude, cette fois-ci simplement sur l'appareil lave-vaisselle. (cf. *Annexes : Figure 23, Tableau 12 et Tableau 13 page 57 à 56*)

On trouve une certaine cohérence notamment l'absence d'utilisation durant la nuit (de 00:00 à 08:30, étendue jusqu'à 10:00). En revanche les durées des plages retenues sont curieuses en comparaison avec un cycle standard qu'on dirait être plutôt aux alentours de 30 min (on en est loin...). On observe que la fréquence la plus élevée d'utilisation est en début d'après-midi (de 12:30 à 14:30), ce qui peut paraître cohérent par exemple dans le scénario suivant : *après avoir mangé, l'habitant lance le lave-vaisselle.*

On réalise une quatrième étude, cette fois-ci simplement sur la lumière de la cuisinière.

A partir des figures disponibles en *Annexes (Figure 24, Tableau 14 et Tableau 15 page 59 et 60)*, on fait les remarques suivantes.

On observe des blocs dans l'utilisation de la lumière ou plutôt concernant l'enclenchement de cette activité :

- Le matin de 5 :00 à 07 :00
- Le soir de 17 :30 à 20 :30 (mais aussi sur le coup de 21 :30-22 :00 mais un peu moins)
- On n'observe pas de bloc distinct concernant l'horaire du midi. Cela peut vouloir dire que l'habitant ne mange peut-être pas tous les midis chez lui, ou bien que son horaire du repas du midi est trop étalé et ne possède pas de créneau donc d'habitude précise.

#### ***2.3.4.5 Activité de la salle de bain***

Utilisation des capteurs suivants :

- Courant de la machine à laver (SMP05 - 6636)
- Taux d'humidité de la pièce (TH01 - 6222)
- Température de la pièce (TH01 - 6623).
- Lumière de la salle de bain (L03 - 7125)

On finit par étudier ici simplement dans un premier temps la machine à laver puis la lumière de la salle de bain.



## Résultats :

On étudie l'activité de la machine à laver à partir du 2020-03-24. On observe (cf. *Figure 25 en Annexes page 61*) que l'activité est réalisée principalement en journée avec un pic de fréquence d'utilisation pour le début de l'activité à 11:30 (5) [soit 13,5%], et à 15:30 / 18:30 / 13:00 (3) [soit 24,3%]. Ce que l'on remarque surtout c'est une absence d'activité la nuit de 21:30 à 09:00 (aucune activité) [soit 50% de créneaux vides], plage étendue à partir de 19:30 (aucune ou très faible activité) [soit 58,3% de créneaux vides]. Cette activité ne peut être exercée que sur une moitié précise de la journée.

D'après la figure en *Annexes (Tableau 16 page 61)*, on peut poser un commentaire critique sur la durée du cycle de machine qui paraît court (entre 15 et 25 minutes généralement).

Lors de l'étude sur la semaine (*Tableau 17 page 62*), le nombre faible de plages à l'étude (38) ne permet pas de dégager de grande tendance concernant "le jour de lessive". Néanmoins, sur ces 38 plages, il semble que ce soit le lundi sur l'heure du midi qui l'emporte. Une plus grande quantité de données pourrait confirmer ou infirmer cette assertion.

Concernant la lumière de la salle de bain (cf. *Figure 26 en Annexes page 63*) :

On étudie les données de la lumière à partir du 2020-02-26 pour un total de 246 plages. On remarque qu'on a de plus fortes activités en soirée. On retient les pics de valeurs suivants :

- Pic le plus important à 22 :30 (15) [soit 6,1%]
- Autres pics importants à 12 :00 / 21 :30 / 23 :30 (11) [soit 13,4%], à 13 :00 / 17 :00 / 00 :30 (10) [soit 12,2%] et à 11 :00 / 17 :30 / 19 :30 / 21 :00 (9) [soit 14,6%].

Ce qu'on remarque surtout c'est l'absence d'activité sur la plage de 04 :30-07 :30 [soit 14,5% de créneaux vides], étendue à partir de 03 :00 [soit 20,8% de créneaux vides].

Avec le tableau en *Annexes (Tableau 18 page 63)*, on remarque des utilisations la nuit, ce qui peut impliquer des réveils la nuit pour aller à la salle de bain par exemple.

Ensuite, à l'aide du *Tableau 19 en Annexes (page 64)*, on remarque que les réveils la nuit qu'on a pu remarquer se déroule de manière plutôt ponctuelle en semaine (après 00 :30), et rare, contrairement au week-end. Cela peut éventuellement coïncider avec des moments où l'habitant se couche tard.

## 2.4 Etablissement de scénarios

Nous avons travaillé l'établissement de scénario par la méthode des chaînes de Markov en menant plusieurs analyses. Dans un premier temps, j'ai réalisé une modélisation simpliste en créant une liste ordonnée chronologiquement des capteurs dits « en activité », c'est-à-dire lorsque le capteur est considéré en état de marche, en s'appuyant sur les plages créées précédemment mais en faisant fi des durées de leur état. Différentes analyses ont été menées notamment en faisant varier les capteurs utilisés pour cette analyse. Dans un deuxième temps, j'ai souhaité améliorer cette modélisation en proposant un modèle de 3 chaînes regroupant différents capteurs dont ces derniers pouvaient avoir des activités « en marche » simultanées. Ces différentes modélisations ont été réalisées pour l'ensemble des données disponibles, celles-ci extraites à la minute, et les démarches seront explicitées plus précisément dans leurs parties correspondantes ci-dessous.

### 2.4.1 Première modélisation dite « simple »

Pour cette première étude, la modélisation est réalisée sous forme d'une chaîne de caractère unique où les activités sont ordonnées chronologiquement, ces dernières étant inscrites à cette chaîne lorsqu'elles sont considérées dans un état « en marche », comme par exemple, « la lumière est allumée » ou bien « une présence est détectée dans la pièce x ». Pour cela, nous représentons chaque activité par un caractère prédéfini de la manière suivante :

# a : entrer dans la salle de bain  
# b : entrer dans la livingroom  
# c : entrer dans la cuisine  
# d : entrer dans le couloir  
# e : entrer dans la chambre  
# f : ouvrir frigo  
# g : ouvrir porte balcon  
# h : ouvrir porte entrée  
# i : dormir  
# j : regarder la télé  
# k : lumière de la cuisinière  
# l : lave-vaisselle en marche  
# m : micro-onde en marche  
# n : machine à café en marche  
# o : sandwichmaker en marche  
# p : machine à laver en marche  
Soit 16 capteurs

Précisément, cette chaîne relate les activités par rapport au début de celle-ci, concernant des activités relevées toutes les minutes et réunies sous forme de pages. Les activités résultent généralement de l'activité d'un unique capteur, hormis par exemple l'activité « i : dormir » qui est la combinaison de deux capteurs comme cela a été étudié dans la première partie des réalisations de ce rapport.

Une première sous-modélisation consiste à former des blocs d'activités, un bloc concernant des activités qui démarrent simultanément, qui sont séparés par des virgules. Un bloc peut contenir une unique activité, c'est-à-dire qu'aucune autre activité n'a été trouvée en état de marche simultanément selon l'horaire de début de l'activité. On intitulera alors ces blocs « événements ». Avec cette modélisation, on cherche à connaître la probabilité des caractères suivants un caractère choisie, en tenant compte alors de ne pas considérer une activité qui surviendrait de manière simultanée.

On peut déjà porter une première critique à cette modélisation sur le fait qu'on ne prend pas en compte la durée des activités et ne prend en compte les activités dites simultanées que si celles-ci *débutent* en même temps, mais non si celles-ci interviennent en même temps *sur la durée*.

Ensuite, une seconde sous-modélisation consiste à ne pas former ces blocs. Il s'agit donc d'une chaîne de caractère sans virgule de séparation, mais pour laquelle j'ai mené d'autres analyses. Il s'agissait ici de reconnaître quelles séquences de caractères revenaient le plus souvent.

Une nouvelle fois, cette modélisation peut être critiquée sur le fait qu'on ne considère pas si les activités se déroulent « en simultanées » (comme défini par la modélisation précédente), et sinon lorsque celles-ci sont simultanées, la séquence d'activités simultanées et alors ordonnées de manière alphabétique.

Ainsi, présenté ci-dessous les parties 2.4.1.1 à 2.4.1.4 concerne cette première sous-modélisation (c'est-à-dire celle avec des blocs séparés par des virgules) et les parties 2.4.1.5 et 2.4.1.6 sont des analyses réalisées avec la seconde sous-modélisation (c'est-à-dire la chaîne de caractères sans les virgules).

#### *2.4.1.1 Analyse sur l'apparition de la lettre suivante*

Premièrement, avec cette première sous-modélisation nous obtenons une chaîne de 33 170 caractères, où sont comptabilisées les virgules, soit 18 544 activités (caractères sans virgule dans la chaîne) et 14 627 événements (nombre de blocs séparés par des virgules).

Un premier traitement a été réalisé sur la chaîne notamment en supprimant les doublons de caractères pour des activités uniques. Après réflexion, nous réaliserons de nouvelles analyses avec un second traitement de la chaîne qui supprimera aussi les doublons de blocs de caractères, l'idée étant de connaître qu'elle est la *nouvelle* activité qui survient, supposée donc différente de la précédente.

Dans un premier temps, j'ai regardé l'occurrence des caractères au sein de la chaîne, dont les résultats sont réunis dans le *Tableau 20 en Annexes (page 65)*.

Avec ce tableau, on remarque que les pourcentages les plus élevés concernent les appareils de détection de mouvement des pièces (c'est-à-dire les lettres : a,b,c,d,e). Néanmoins, on note des présences plutôt faibles pour certaines pièces telles que le salon (b). Ces pourcentages sont tout de même à utiliser avec parcimonie étant donné la qualité et la précision de tels capteurs où plusieurs données non viables ont été relevées précédemment et qui dans cette première analyse n'ont pas été retirées (exemple de la salle de bain à partir de la mi-mars). On peut remarquer que le fort pourcentage de cette pièce dépeint l'anomalie. De plus, certains capteurs ont été installés plus tard ou sur des courtes périodes sur la durée des 6 mois, ce qui entraîne bien sûr une collecte de données plus faible et biaise aussi les résultats.

Ensuite, j'ai réévalué ces occurrences en supprimant aussi les doublons de blocs de caractères comme présenté par le *Tableau 21 en Annexes (page 65)*.

On remarque alors que les doublons de blocs concernent majoritairement les capteurs de mouvement. Il n'y a plus que 13 535 caractères (virgules non comptabilisées, pour une chaîne de 29 665 caractères avec virgules). On observe particulièrement pour certains caractères que malgré la diminution de l'occurrence, il reste majoritaire, voire leurs pourcentages ont fortement augmenté, comme c'est le cas de la salle de bain (a) et de la chambre (e) pour les capteurs de mouvement. On peut dire dès à présent qu'avec ce type de modélisation, et dans le jeu de données même, il existe un biais concernant les données des capteurs de mouvement qui représente un grand volume de données. Cela pose des questions sur la qualité du capteur ou bien la manière dont la donnée a été recopiée toutes les minutes selon le protocole de la réalisation de ce jeu de données. Ces observations m'ont alors conduites à porter de nouvelles analyses en faisant fi de ces capteurs de mouvements, présentées dans les parties 2.4.1.3 et 2.4.1.6.

Après s'être intéressé à la composition de la chaîne, je me suis intéressée aux probabilités d'apparition de la lettre suivante pour une lettre choisie. La lettre choisie ici constitue un bloc à elle seule, c'est donc un événement, tandis que la lettre suivante peut être un événement à elle seule ou appartenir à un bloc d'activités simultanées qui forment un événement. Dans ce dernier cas, il s'agit de la première lettre qui apparaît dans ce bloc.

Le *Tableau 22* (cf. *Annexes page 65*) présentent alors les résultats de probabilités précises au millième et pour une chaîne de caractère qui a été traitée par la suppression de doublons d'activités uniques et de blocs.

On remarque qu'à travers ce tableau, les activités suivantes les plus communes concernent la présence dans les pièces, ce qui n'est pas étonnant suite à l'analyse réalisée précédemment sur l'occurrence des caractères dans la chaîne.

On peut noter aussi une certaine incohérence par exemple dans la succession de certaines pièces. Alors que le couloir (d) est un passage obligé entre celles-ci, il n'est pas majoritaire. On note encore l'activité étrange dans certaines pièces pour certains autres capteurs ne se trouvant pas dans ces pièces, tel que la porte du balcon (g). A contrario, on peut relever une certaine cohérence de l'activité de certains appareils suivit de la présence détectée dans la pièce telle que le frigo (f), le micro-onde (m), la machine à café (n) et le sandwichmaker (o) avec la cuisine (c).

On peut critiquer la précision des analyses menées précédemment de par le fait que l'on regarde la première lettre suivant une lettre choisie. Or, si cette lettre appartient à un bloc, bloc composé indifféremment d'activités débutant en même temps, celles-ci sont alors organisées de manières alphabétiques. Il s'agit alors ici de mener une analyse plus précise en ne regardant pas directement la lettre suivante, mais en tenant compte de toutes les activités présentes dans ce bloc d'activités si ce bloc n'est pas une activité unique. Le *Tableau 23* (cf. *Annexes page 66*) présentent les probabilités relevées.

Nous allons alors comparer les *Tableau 22* et *Tableau 23* (cf. *Annexes page 65 et 66*). Tout d'abord, on observe que les activités qui suivent l'activité choisie sont sensiblement les mêmes, à quelques exceptions (cf. surligné orange dans *Tableau 23*), qui sont alors négligées dû à leur proportion très faible.

Ensuite, nous regardons l'évolution des probabilités entre les tableaux. La plupart de celles-ci diminuent légèrement ou restent sensiblement les mêmes. Quelques proportions fluctuent néanmoins avec de grands écarts tels que pour l'activité e à la suite de l'activité a qui diminue de plus de 0,2 mais reste majoritaire pour l'activité a tout de même (a suivie de e). Certains à l'inverse ont leur proportion qui augmente beaucoup telles que pour l'activité e à la suite de l'activité i.

Enfin, en regardant les activités dominantes, c'est-à-dire l'activité suivante qui a la plus haute proportion après l'activité choisie, les résultats restent sensiblement les mêmes pour les deux tableaux, hormis pour les deux activités b, c et p.

Cette analyse permet surtout de remettre à jour les probabilités plus proches de la vérité, sans le biais de l'ordination alphabétique. Aussi, nous retrouvons de nouveau les biais que nous avons pu constater précédemment notamment sur les capteurs de mouvement ou la cohérence des résultats.

De plus, cette étude comporte une certaine limite du fait de l'échantillonnage des données qui est ici réalisée à la minute. On peut éventuellement imaginer que certains capteurs dont l'activité peut être très brève, tels que les capteurs de contact, que la donnée ne soit pas remontée dans cette chaîne. Il pourrait être intéressant de réaliser des analyses

pour une fréquence d'échantillonnage plus élevées. Néanmoins, dans ce désir de précision, nous risquons d'augmenter le biais du volume de données des capteurs de présence évoqué précédemment. Les nouvelles données précieuses que nous tentons d'extraire seraient toujours noyées par le volume des données des capteurs de mouvement. Une solution envisageable serait alors de réaliser cette étude avec une fréquence plus élevée, mais alors sur une plage de données restreinte et définie, qui réunit l'ensemble des appareils. Une difficulté est posée ici puisqu'il faut pouvoir faire la réunion des plages des appareils pour obtenir cette dite plage d'étude mais aussi tenir compte des plages pour lesquelles nous avons observés des comportements anormaux de la part des capteurs de mouvement (ex. de la salle de bain citée précédemment). Mon camarade Abderrhamane s'est penché sur cette étude.

#### *2.4.1.2 Analyse sur l'appartenance d'activités à un bloc*

En reprenant le protocole, puisque l'appartement doit être occupé par un unique habitant, il peut paraître suspicieux d'avoir des activités qui *débutent* simultanément. On se rend compte que cela est éventuellement possible pour certaines combinaisons de capteurs particulières, telles que des capteurs situés dans la même pièce, mais aussi selon leur type, par exemple des capteurs de présence avec des capteurs liés à des appareils électroniques. Il s'agit alors de vérifier ceci ici.

Le *Tableau 24* (cf. *Annexes page 66*) présente alors les résultats d'appartenance d'une lettre dans un même bloc qu'une lettre définie. On remarque principalement que la plupart des activités sont en réalité unique avec en moyenne une proportion de 0,76. Ainsi, pour une lettre choisie, cette activité est unique les  $\frac{3}{4}$  du temps et appartient à un bloc pour  $\frac{1}{4}$ .

On note aussi une coexistence avec une proportion entre  $\frac{1}{4}$  et  $\frac{1}{3}$  pour les lettres a et e. On soutient que alors que ce phénomène est dû à la disproportion des données sur les différents capteurs, vu précédemment. Cela renforce l'idée de réaliser une étude sans les capteurs de mouvement pour étudier plutôt le comportement des appareils.

#### *2.4.1.3 Analyse sur l'apparition de la lettre suivante sans les capteurs de mouvement*

Comme vu précédemment, la proportion des données selon le type de capteurs semble induire un biais dans les analyses et ne nous permet pas de voir distinctement l'enchaînement d'utilisation entre les différents appareils. Il s'agit alors ici d'appliquer les mêmes méthodes mais en retirant les capteurs de mouvement de l'étude. On réutilise donc les mêmes lettres (de f à p, ce qui correspond à 11 capteurs). On réalise l'analyse sur une chaîne traitée sans doublons d'activités uniques ou de blocs d'activités et pour les activités suivantes qui appartiendraient à tout le bloc suivant, sans distinction de sa position dans ce bloc.

On obtient une chaîne de 1942 caractères avec virgules (découpage en blocs), soit 961 caractères (virgules non comptabilisées). On voit une différence de proportion significative en comparaison aux valeurs précédentes avec l'utilisation des capteurs de mouvement. Celle-ci est flagrante alors que les capteurs de mouvement ne représentaient que moins d'un tiers pour l'ensemble des capteurs (c'est-à-dire 5 capteurs de mouvement pour une totalité de 16 capteurs).

J'ai réévalué la proportion des termes dans la chaîne, présentée par le *Tableau 25* (cf. *Annexes page 67*). On observe une nouvelle fois des proportions disparates selon les capteurs partant de 2% jusque 23%. Ceci est à mettre en contraste selon le temps d'installation des capteurs qui n'est pas le même pour chacun et qui induit des biais dans nos analyses sur la totalité des données. Aussi, une remarque que l'on peut faire sur la chaîne obtenue est que celle-ci contient généralement des blocs d'activités uniques (c'est-à-dire qu'il n'y a presque plus d'activité simultanées).

J'ai donc réuni au sein d'un tableau (*Tableau 26 en Annexes page 67*), les probabilités d'apparition de la lettre suivante comme cela a été fait dans les analyses précédentes.

Nous allons comparer ces résultats aux résultats précédents présentés par le *Tableau 23* (*page 66*). Tout d'abord, on remarque que les activités suivantes ne suivent jamais une activité choisie :

- Les activités ne se suivent jamais elles-mêmes sauf pour les activités g et k.
- L'activité h ne suit jamais les activités m et o.
- L'activité j ne suit jamais les activités m, o et p.
- L'activité l ne suit jamais les activités m et p.
- L'activité m ne suit jamais les activités j, l et o.
- L'activité n ne suit jamais l'activité j.
- L'activité o ne suit jamais les activités j, m et p.
- L'activité p ne suit jamais les activités l, m et n.

On remarque que de nouvelles séquences d'activités ont été ajoutées, ceci bien sûr dû à la suppression des capteurs de mouvement de l'analyse. Il est difficile de comparer les maximums des parties de tableaux concernées, considérant les proportions très différentes, les maximums de la partie du *Tableau 23* étant généralement de l'ordre du 100<sup>ème</sup>. Néanmoins, si on se prête à cette comparaison, on peut dire que l'on retrouve vraisemblablement les mêmes maximums.

Maintenant, en analysant le *Tableau 26* seul, et en regardant les activités les plus importantes qui se suivent, on remarque les choses suivantes :

- On retrouve l'activité « dormir » (i) qui suit « regarder la télé » (j), ceci probablement le soir au coucher.
- Ensuite, l'activité « ouvrir le frigo » (f) qui suit « allumer la cuisinière » (k).
- Les activités « ouvrir le frigo » (f) et « allumer la cuisinière » (k) qui suivent « mettre le micro-ondes en marche » (m) ou « mettre le sandwichmaker en marche » (o).

Pour les autres activités, leur séquence paraît moins intuitive. On remarque notamment les activités d'ouverture de porte (f), (g) et (h) qui se suivent mutuellement ; ou encore aller dormir (i) après avoir allumé la cuisinière (k).

Ces résultats peuvent une nouvelle fois être remis en cause par la variabilité des proportions des données très différentes selon les capteurs. Aussi, une analyse plus fine avec une fréquence d'échantillonnage plus élevée sur une période précise pourrait venir confirmer ou infirmer ces résultats, étude sur laquelle mon camarade de stage s'est penché.

#### 2.4.1.4 Analyse sur l'apparition de la lettre suivante avec un changement d'activité

Compte-tenu de la disproportion des données des capteurs, on cherche une nouvelle fois ici à diminuer ce biais en regardant l'activité « faire un repas » (q) qui réunit les capteurs frigo (f), micro-ondes (m) et sandwichmaker (o). On obtient alors une chaîne de 1883 caractères (virgules comptabilisées), soit 933 caractères (virgules non comptabilisées), avec très peu de blocs d'activités simultanées mais plutôt des activités uniques. Cette dernière remarque n'est pas surprenante puisque c'était déjà le cas lors de la précédente analyse. Le *Tableau 27* (cf. *Annexes page 67*) fait état des occurrences des termes au sein de la chaîne.

J'ai mené alors les mêmes analyses que précédemment, dont les résultats sont réunis dans le *Tableau 28* (cf. *Annexes page 68*). Nous allons comparer celui-ci au tableau précédent (*Tableau 26 en Annexes page 67*).

On retrouve vraisemblablement des valeurs similaires pour les activités communes, ainsi que les maximums similaires, hormis le fait que c'est principalement l'activité « faire un repas » (q) qui domine à présent, ce qui se témoigne par sa proportion dans le *Tableau 27*. Cela est notamment dû au capteur du frigo qui se trouvait majoritaire parmi les trois capteurs réunis dans les études précédentes.

A la suite de ces différentes analyses menées, on se rend compte des limites de cette modélisation globale et notamment des biais au sein du jeu de données, notamment pour les capteurs de type détection de mouvement. Celles-ci présentent un volume de données très important, ce qui est sans surprise, mais ce qui nous empêche de faire des analyses pour des données de volume plus faible. A terme, dans l'objectif d'implémenter une intelligence artificielle à partir de ces données, puisque l'apprentissage automatique repose principalement sur le volume de données fourni, les données de certains appareils en plus faible quantité pourrait être négligées par les algorithmes et empêcher des prédictions pertinentes. Ces différentes analyses ont alors permis de mettre en lumière cette difficulté et suggère alors de réserver un traitement à part pour les capteurs de type de détection de présence.

#### 2.4.1.5 Analyse d'occurrence de séquences de caractères

Pour rappel, nous utilisons une modélisation qui ne sépare pas les événements, c'est-à-dire qui ne contient pas de virgule donc pas de blocs. Il n'y a donc aucune distinction des événements dans la chaîne si certains sont simultanés ou non. Le traitement appliquait à la chaîne est uniquement la suppression des doublons de lettres qui se suivent et non de groupes de lettres comme cela a pu être fait pour d'autres analyses.

Ici, nous allons chercher à déterminer en choisissant en paramètre la longueur de la séquence de caractères, quelles sont les séquences qui reviennent le plus avec leur probabilité d'apparition dans toute la chaîne. Nous réalisons cette probabilité selon leur proportion dans le texte, c'est-à-dire par rapport au nombre de caractères total et non au nombre de combinaison de chaînes possibles.

Dans un premier temps, on a regardé les combinaisons pour selon une première chaîne *kgram1*, la lettre suivante pour former *kgram2* était celle ayant la plus forte probabilité de se trouver à sa suite. Dans un second temps, on a amélioré le code de sorte qu'on obtienne toutes les combinaisons possibles (c'est-à-dire pas de sélection selon la plus forte probabilité

pour la lettre suivante). Bien sûr, si la lettre suivante a plus de probabilité d'apparaître que d'autres lettres, alors la nouvelle combinaison a une plus forte probabilité que pour les autres combinaisons. Mais si l'on souhaite comparer des combinaisons de différentes tailles, alors on peut avoir des probabilités équivalentes ou supérieures pour des mots d'une taille  $n$  par rapport à des mots de taille  $n-1$ , tandis que la propriété suivante est vraie : plus le mot est petit, plus sa probabilité d'apparition dans le texte peut être grande (*a contrario*, plus le mot est grand, plus sa probabilité d'apparition dans le texte est faible). De ce fait, j'ai regardé les séquences jusqu'à 6 caractères inclus.

Ainsi, j'ai cherché à identifier s'il existe des enchaînements d'activités qui formeraient un pattern particulier. Avec pas plus de 6 caractères, j'ai souhaité rester sur des séquences simples. De plus, j'ai recensé les expressions qui auront une probabilité de plus de  $0.5 \times 10^{-2}$  d'apparition dans le texte, négligeant les autres jugées non significatives. Il s'agit d'une valeur seuil pour faire ressortir le plus de combinaisons possibles, au sein desquelles j'ai cherché à identifier quelles séquences semblent être significatives.

Dans un premier tableau (*Tableau 29 en Annexes page 68 à 69*), est répertorié ces expressions ordonnées selon leur probabilité d'apparition de la plus grande à la plus faible. Comme on a pu le faire remarquer auparavant, nous allons retrouver beaucoup de séquences avec les lettres correspondant aux capteurs de mouvement, ces derniers ayant un volume de données plus conséquent. On peut remarquer que nous répertorions aussi beaucoup de séquences qui ne semblent pas porter de sens *a priori*, l'exemple principale étant des mouvements entre les pièces sans la détection de présence par le couloir (d) qui est un passage obligé. Si l'on regarde la configuration de celle-ci et l'emplacement du capteur, cette donnée *ne devrait pas* manquer. On s'interroge donc sur la qualité du capteur ou bien la remontée des données de celui-ci qui ne semble alors pas fiable. De manière générale, à la suite des observations réalisées par ce type de capteurs, ces données sont plutôt difficiles à traiter, de part leur volume mais aussi ce qui semble être une défaillance. On peut supposer que ces défauts sont récurrents pour ce type de capteurs et qu'il faudra alors porter un traitement particulier à celle-ci pour réaliser de l'apprentissage par exemple, comme dit précédemment. Une nouvelle fois, j'ai mené alors une autre analyse sans les données des capteurs de mouvement.

#### *2.4.1.6 Analyse d'occurrence de séquences de caractères sans les capteurs de mouvement*

Dans cette partie j'ai repris la même méthodologie que précédemment, mais j'ai réalisé ici l'étude sans les capteurs de mouvement, c'est-à-dire que leurs caractères associés (a à e) ne sont pas présents dans la chaîne.

Nous retrouvons les résultats en *Annexes* dans le *Tableau 30 (page 69 à 70)*. Contrairement au tableau précédent, il est difficile de parler de combinaisons insensées puisque les séquences évoquent simplement la mise en marche d'appareil. Mais, nous pouvons remarquer que nous avons des séquences, souvent celles présentant le pourcentage le plus élevé pour une longueur de séquence donnée, qui ont des caractères très répétitifs du type « xyxyxy », ceci décrit à partir d'une séquence de taille 3. Ceci peut décrire un problème au niveau de la qualité du capteur ou de la remontée de la donnée, ce qui contribue au biais que nous pourrions avoir pour une implémentation plus poussée. On notera aussi que même pour les expressions de petite taille (la plus faible étant de 2), le pourcentage d'apparition ne dépasse pas les 10%. Il est donc difficile de décrire une généralité, même si l'on peut distinguer



certaines activités plus récurrentes. Cela est dû à un plus grand échantillon de séquences pour cette étude ici : on compte 43 expressions de taille 2 pour ce tableau-ci contre 24 pour le tableau précédent, soit près du double. Aussi, les activités (ou appareils) qui semblent revenir en tête sont : la porte du frigo, la porte du balcon, dormir, la lumière de la cuisinière, ainsi que la porte d'entrée.

#### 2.4.2 Deuxième modélisation à 3 chaînes

Ici, j'ai cherché à revoir la modélisation de la chaîne de caractère pour déceler les séquences d'activité et les incohérences éventuelles de ces séquences. Cette modélisation consiste en plusieurs chaînes de caractères dont les activités séparées par des virgules correspondent à une tranche temporelle précise. Les trois chaînes correspondent à :

- Une chaîne de « contexte » qui relate les données des capteurs de mouvement et de porte, c'est-à-dire la lettre indique la présence de l'individu dans une pièce *i*. Si plusieurs pièces sont identifiées pendant cette tranche temporelle, on a alors une *incohérence* (par exemple, un non-respect du protocole, un défaut du capteur ou de la remontée de la donnée).
- Une chaîne « principale » qui relate la séquence d'activité considérée comme activité principale (donc une unique lettre), c'est-à-dire elle relate l'activité qui se déroule à un instant *t*, qui est la seule activité à ce moment-là ou l'activité la plus longue à partir de cet instant. Le choix de cette activité détermine la tranche temporelle utilisée pour les deux autres chaînes.
- Une chaîne « secondaire » des activités dites secondaires, c'est-à-dire qui se déroulent durant la tranche d'activité principale. Il peut s'agir d'une activité unique qui ne dure pas forcément la même durée que l'activité principale mais qui se déroule pendant cette durée, ou bien une séquence de plusieurs activités.

Lorsqu'aucune activité n'est trouvée, on remplit la chaîne par un 'tiret' et on passe à l'instant présent. Cela pourra nous permettre par exemple de trouver les instants d'inactivité dans l'appartement.

Dans le scénario où l'activité commencerait pendant une activité principale *a1* et se terminerait pendant une activité *a2*, on considérera pour cette modélisation l'existence de cette activité uniquement en simultanée pendant l'activité *a1*. Cela n'a pas pu être réalisé, mais il pourrait être intéressant de vérifier que la durée des activités secondaires est bien comprise pendant la durée de l'activité principale et adapter la modélisation pour ce cas particulier.

Avec les différents codes établis, notamment la fonction *get\_info\_chaine*, on obtient alors pour un instant *t* précis de la forme 'AAAA-MM-DD HH:mm:ss' toutes les informations pour chaque donnée pour les trois chaînes telles que : l'indice du capteur correspondant dans la liste des capteurs utilisés (donc indication sur le capteur « en marche »), ligne de la donnée récupérée correspondant au tableau du capteur, instant *t* choisi en entrée pour trouver l'activité correspondante, date et horaire du début de l'activité récupérée, date et horaire de la fin de l'activité récupérée. Toutes ces informations permettent alors de voir précisément l'enchaînement de l'utilisation des différents appareils et notamment de déceler avec précision si certains appareils sont enclenchés en simultané ou non, ce qui pourrait indiquer la présence de plus d'une personne dans le logement (et un non-respect du protocole annoncé).

Les résultats ont été réunis dans un tableau de 77455 lignes, enregistré sous format Excel afin de pouvoir être réutilisé facilement par l'équipe qui travaille toujours sur le projet. J'ai réalisé différentes analyses pour interpréter ce tableau :

On compte 67 028 lignes qui comportent des '-' pour les trois chaînes, c'est-à-dire qu'aucune donnée de présence (contexte) ou d'activité (principale ou secondaire) a été relevée. A noter que la remontée des données s'arrête aux alentours du 2020-08-26 et que le tableau s'arrête lui un jour après soit le 2020-08-27 au soir, donc nous avons au minima une journée « vide ». Il avait été choisi de garder une marge pour ne pas manquer de données.

Cela représente 86,5% du tableau (pour des données échantillonnées à la minute). On comprend alors que sur toute la durée de l'expérience, il peut y avoir des grandes périodes de non-activité, plus ou moins dispersées. A noter que pour des données échantillonnées à la minute, cela aurait été de l'ordre de 260 000 lignes. Or ceci est réduit de 70% avec cette modélisation qui cherche les activités pour la tranche temporelle de l'activité principale et met alors en contraste le pourcentage élevé de créneaux vides.

Parmi les 67 028 lignes, 42 731 lignes ont été identifiées selon le filtre suivant : « On compte les tirets si et seulement si, pour la présence détectée avant le bloc de tirets, la présence détectée après ce bloc est différente de la précédente, et il n'y a aucune activité détectée ».

En effet, dans les conditions où une même lettre indiquant la présence dans une pièce est séparée par un bloc de tirets, on peut supposer que la personne est restée dans la pièce mais qu'alors soit elle ne bougeait plus, soit la donnée de sa présence n'est pas remontée. Ainsi, les 42 731 lignes correspondent au nombre de tiret pour des successions de pièces différentes, soit plus « réellement » 55% de créneaux vides.

Maintenant, on cherche à savoir parmi ces lignes, si l'enchaînement des pièces est correct, c'est-à-dire est-ce que la présence a été détectée dans le couloir, passage obligé d'une pièce à l'autre. On relate 6 108 lignes pour lesquelles le bloc de tirets est suivi par une présence détectée dans le couloir. On a donc 36 623 lignes soit 47% de la chaîne de présence qui relatent « d'incohérences » comme évoqué plus tôt. De plus, on calcule en moyenne que les blocs de tirets successifs sont de 23 lignes, c'est-à-dire au moins 22 minutes.

Ensuite, 2 174 lignes ont été identifiées comme ayant des chaînes de contexte de plus d'un élément strictement, c'est-à-dire qu'il a été identifié une présence simultanée dans les pièces dans le cas si aucune activité principale a été détectée, ou, si une activité principale a été détectée, cela peut être une succession de présence dans différentes pièces pendant la durée de cette activité principale. Parmi elles, on compte alors 1 611 lignes qui concernent des présences simultanées pour lesquelles aucune activité n'a été relevée (donc en à la même minute). En regardant le tableau de résultats, il ne s'agit pas de doublons et celles-ci semblent de longueur 2 ou 3 caractères au plus.

De plus, on relate 1 150 lignes pour lesquelles on détecte une activité principale (possibilité d'activité secondaire ou non), ce qui représente 1,5% de la chaîne. Cela se comprend du fait que ce sont bien les données concernant les capteurs de mouvement qui sont majoritaires et non les autres appareils.

# 3 Conclusion

## 3.1 Retour d'expérience

J'ai beaucoup aimé réaliser ce stage et plus particulièrement travailler sur ce sujet des « smart homes » qui est un domaine qui m'intéresse. Cela m'a permis d'en savoir plus sur le sujet et à travers cet aspect de la recherche de connaître les états d'avancement dans le domaine et les manières de procéder.

Cette expérience m'a donné aussi une vision plus globale de la recherche en étant imprégné dans celui-ci de manière continue. Cela m'a fait remettre en perspective certains attendus que nous pouvons avoir lorsque nous cherchons quelque chose en fonctionnant sur nos a priori, mais surtout la difficulté lorsque l'on se confronte à la réalité de ce que les études peuvent nous apporter. La difficulté est alors aussi de démêler le lot d'informations que nous pouvons récolter et de tirer de là une généralité. Cela a été un vrai challenge lors de ce stage, notamment du fait des biais que nous pouvions trouver dans nos données.

J'ai pris aussi alors conscience de la frustration que nous pouvions ressentir en tant que chercheur avec ces difficultés, notamment sur les limitations que nous pouvions avoir sur les données ou les biais nombreux que nous pouvions recenser. Nous pouvons avoir une vision inspirante du projet mais nous réalisons qu'il y a encore beaucoup de travail et de recherche avant d'atteindre celle-ci.

J'ai apprécié réaliser ce stage en équipe, surtout puisque nous étions en autonomie. Cela nous a permis d'échanger et de faire part de nos difficultés pour éventuellement trouver de nouvelles perspectives lorsque la frustration était trop grande. Le fait d'être en autonomie pour la plupart de ce stage pouvait être parfois déconcertant dans ces cas-là, mais j'ai trouvé aussi cela formateur en ce sens. Cela m'a confirmé aussi qu'il est bien mieux, selon moi, de partager son projet (quel qu'il soit) et de travailler en équipe, ne serait-ce que pour se confronter à une vision différente et ce qui peut déboucher sur de nouvelles idées et solutions.

A l'heure actuelle, je ne pense pas poursuivre dans le domaine de la recherche, mais je ressors enrichie de cette expérience.

## 3.2 Suite du projet

En parallèle des études que nous avons réalisées, le doctorant **Q. PEREZ** a construit un mapping de l'appartement étudié, en lien avec l'objectif des jumeaux numériques que nous n'avons pas pu approfondir. Nous pensons que la dernière modélisation réalisée (celle des trois chaînes) pourrait être alors un bon support pour observer les enchaînements des enclenchements des différents appareils et des présences dans les pièces. Cela permettra de voir concrètement les anomalies et ensuite de les étudier plus précisément avec les informations précises que nous pouvons tirer de ces chaînes comme évoqué précédemment.

Aussi, il s'agit de poursuivre le travail vers les objectifs initiaux annoncés, et notamment de voir de plus amples analyses et traitements concernant les données des capteurs de mouvement qui induisent un biais dans le jeu de données, comme nous avons pu le voir au travers de ce rapport.

# Annexes

## Description des rôles des différentes parties prenantes du projet

Ceci fait référence à la partie suivante du rapport : *1.3 Présentation de l'entreprise*

Différentes organisations participent au projet HUT et ont des rôles qui leurs sont propres à chacun :

- IMT Mines Alès : expertise dans les systèmes de captation de données.
- DMeM : développement d'un « vélo 2.0 » avec pédalage excentrique pour lutter contre la sédentarité
- LiCEM : étudie les questions de collecte et de traitement des données ainsi que la valorisation de leurs usages et les conditions de contractualisation dans l'appartement connecté.
- EPSYLON : évaluation de l'impact d'un environnement hyper-connecté sur la charge cognitive et l'éventuelle atteinte de la mémoire prospective, de la dynamique de la motivation.
- EuroMOV : vise à favoriser la fertilisation croisée de l'intelligence artificielle, des sciences du mouvement et de la santé pour comprendre la plasticité comportementale de l'homme afin d'améliorer sa performance sensomotrice et d'envisager de nouvelles approches thérapeutiques.
- LIRMM : recherche sur la conception et la vérification de systèmes intégrés, mobiles, communicants, et la modélisation de systèmes complexes
- Espaces Dev : avec LIRMM, ils s'interrogent sur l'optimisation de réseaux de capteurs par la question de la classification et de l'orientation des données dans des systèmes informatiques complexes pour un retour plus adapté aux souhaits des utilisateurs.
- IES : analyse des usages des capteurs et systèmes intelligents existants ou à venir et leurs évolutions guidées, en fonctions des défis technologiques et des besoins réels des utilisateurs
- MRM : étudie la création de valeur induite par l'usage des objets, voire leur agencement, la relation homme/objet et leurs influences sur les comportements et le bien-être des habitants.
- LERASS : axes de recherche sur la santé, l'information, les sciences, les savoirs, les pratiques, mes mondes culturels et les expériences numériques.
- Université de Paul Valéry : expertises sur le comportement humains dans les logements et dans la ville, aux dispositifs d'interaction humain/machine ou encore ou déploiement d'interface de médiations innovantes.
- DELIDED : présente sa technologie Maestro qui maîtrise la lumière pour votre bien-être.
- SYNOX : entreprise reconnue dans le déploiement de projets IoT et des réseaux nécessaires à leur connectivité.
- ENEDIS : entreprise spécialisée dans la gestion du réseau de distribution public d'électricité.
- IKEA : permet pour le projet HUT d'utiliser, de partager et d'enrichir les connaissances en matière d'aménagement de la maison pour mieux appréhender la vie à la maison de demain.
- NEXITY : entreprise de promotion immobilière qui a construit le logement étudié dans le projet.

- Oceasoft : expertise dans les études thermiques sur la chaîne du froid et s'investit dans le développement de recherche sur la chaîne du froid avec les usages des consommateurs.
- Université de Montpellier : 8 laboratoires de recherche (santé, informatique, ingénierie, économie, etc.) impliqués.
- Montpellier Méditerranée Métropole :
- CNRS : le projet a pu profiter de l'expertise interdisciplinaire et de ma mise d'outils performants de valorisation.
- MSHSud : attribue un espace de travail au sein de ses locaux pour HUT.

### Plan de l'appartement du jeu de données

Ceci fait référence à la partie suivante du rapport : 2.3.1 *Présentation du jeu de données*

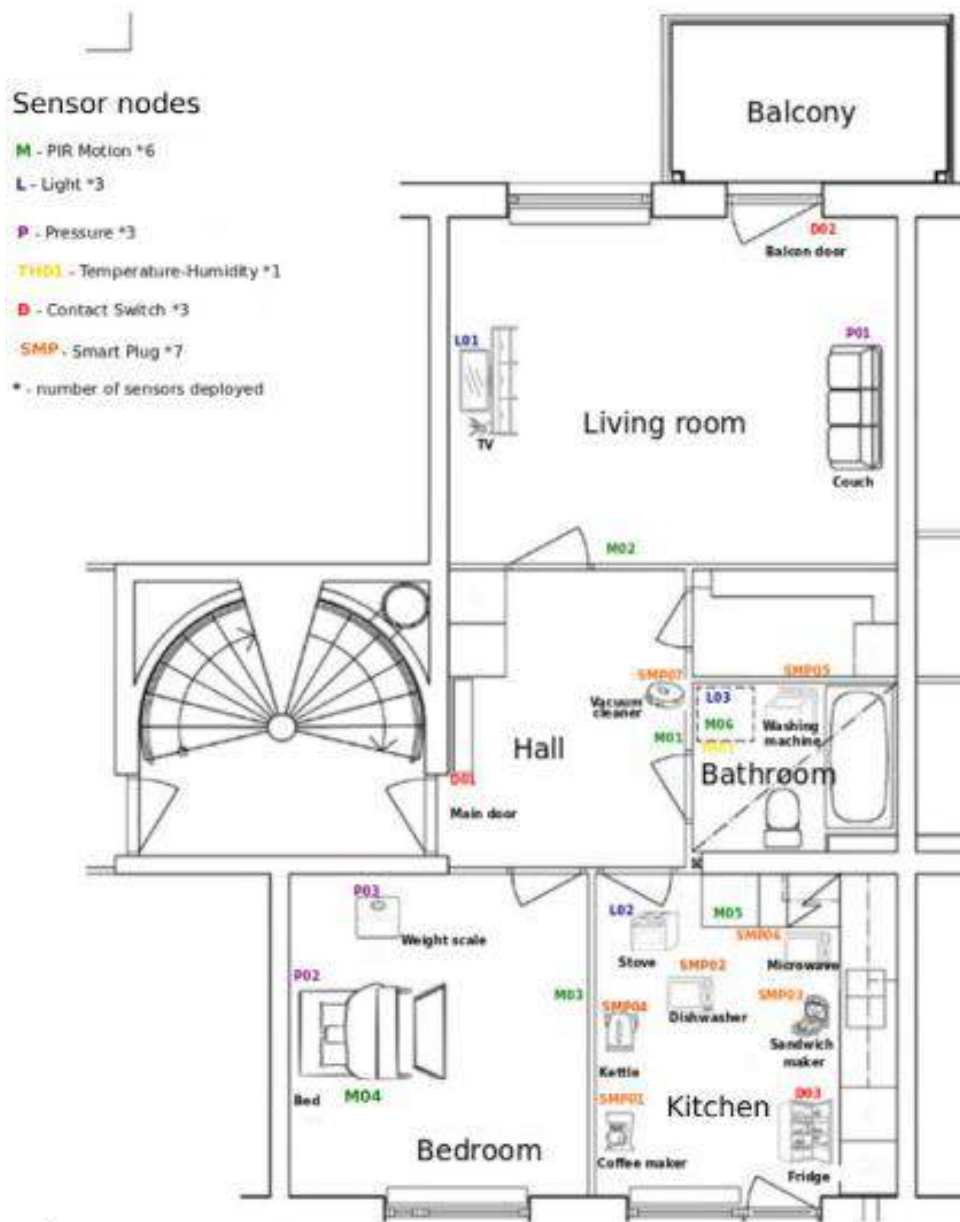


Figure 8 : Plan de l'appartement du jeu de données étudié avec disposition des capteurs

## Tableau récapitulatif des capteurs et de leurs caractéristiques importantes

Ceci fait référence à la partie suivante du rapport : 2.3.3 *Analyse graphique des données*

Légende :

Kitchen		Bedroom		Bathroom		Entrance	
Livingroom		Balcon		Corridor			

Capteur (ID)	ID plan	Type de données	Nom (localisation/objet/ce qu'on mesure)	Plages de valeurs	Etats possibles
5894	M01	INT (binaire)	corridor/38mbiance/motion	De 02-2020 à 08-2020	Présence détectée Aucune présence détectée
5895	M06	INT (binaire)	bathroom/38mbiance/motion	De 02-2020 à 05-2020	Présence détectée Aucune présence détectée
7125	L03	INT (non binaire)	bathroom/38mbiance/light	De 05-2020 à 08-2020	Min<15 : lumière allumée Seuil=1024 : lumière éteinte <b>Quant est-il des entre deux ?</b>
5896	P02	INT (non binaire)	bedroom/bed/pressure	De 02-2020 à 08-2020	>640 : Présence détectée sur le lit <640 : Aucune présence détectée sur le lit <b>Les seuils du lit ne sont pas les mêmes selon la période d'étude</b>
6127	L01	INT (non binaire)	livingroom/tv/light	De 02-2020 à 08-2020	Min<15 : lumière allumée Max=1024 : lumière éteinte <b>Quant est-il des entre deux ?</b>
6220	D02	INT (binaire)	balcon/door/contact	De 02-2020 à 08-2020	Porte ouverte Porte fermée
6253	D03	INT (binaire)	kitchen/fridge/contact	De 02-2020 à 08-2020	1 = Porte ouverte 0 = Porte fermée
6632	SMP01	FLOAT	kitchen/coffeemaker/current	De 03-2020 à 08-2020	Min=0 : machine à café éteinte Seuil=1170 : machine à café allumée
6633	SMP03	FLOAT	kitchen/sandwichmaker/current	De 03-2020 à 08-2020	Min=0 : machine à sandwich éteinte Seuil=1750 : machine à sandwich allumée

6634	SMP02	FLOAT	kitchen/ dishwasher/ current	De 03- 2020 à 08-2020	Min=0 : lave-vaisselle éteinte Seuil=1900 : lave-vaisselle en marche
6635	SMP04	FLOAT	kitchen/ kettle/ current	De 03- 2020 à 08-2020	Min=0 : bouilloire éteinte Seuil=700 : bouilloire en marche
6636	SMP05	FLOAT	bathroom/ washingmachine / current	De 03- 2020 à 08-2020	Min=0 : machine à laver éteinte Seuil>1800 : machine à laver en marche
6896	SMP06	FLOAT	kitchen/ microwave/ current	De 04- 2020 à 08-2020	Min=0 : micro-onde éteint Seuil>1200 : micro-onde en marche
6686	M04	INT (binaire)	bedroom/ ambiance_under_ the_bed/ motion	De 03- 2020 à 08-2020	1 = Présence détectée 0 = Aucune présence détectée
6687	P03	INT (non binaire ?)	bedroom/ weightscale/ pressure	De 03- 2020 à 08-2020	<b>Valeurs contextuelles</b>
7139	SMP07	FLOAT	corridor/ ilifeRobot/ current	De 05- 2020 à 08-2020	Min=0 : aspirateur éteint Seuil>1100 : aspirateur en marche
6222	TH01	FLOAT	bathroom/ ambiance/ humidity	De 02- 2020 à 08-2020	<30% : taux d'humidité très faible 30 à 45% : taux d'humidité correct 45 à 70% : taux d'humidité bon pour une salle de bain >70% : taux d'humidité très élevé Source : <a href="https://www.humidificateur.pro/sante/taux-humidite-ideal#:~:text=Quel%20est%20le%20taux%20d%E2%80%99humidit%C3%A9%20id%C3%A9al%20pour%20votre,d%E2%80%99humidit%C3%A9%20optimal%20%3A%20entre%2050%25%20et%2070%25.%20">https://www.humidificateur.pro/sante/taux-humidite-ideal#:~:text=Quel%20est%20le%20taux%20d%E2%80%99humidit%C3%A9%20id%C3%A9al%20pour%20votre,d%E2%80%99humidit%C3%A9%20optimal%20%3A%20entre%2050%25%20et%2070%25.%20</a>
6223	TH01	FLOAT	bathroom/ ambiance/ temperature	De 02- 2020 à 08-2020	Seuils : < 0°C : glacial 0 - 13°C : froid 13-18°C : frais 18-22°C : bon

					22-27°C : chaud 27-32°C : très chaud >32°C : canicule
5887	L02	INT (non binaire)	kitchen/ stove/ light	De 02- 2020 à 08-2020	<300 : lumière éteinte >300 : lumière allumée
5888	D01	INT (binaire)	entrance/ door/ entrance	De 02- 2020 à 08-2020	1 = Porte ouverte 0 = Porte fermée
5889	P01	INT (non binaire)	livingroom/ couch/ pressure	De 02- 2020 à 08-2020	>300 : Présence détectée sur le canapé <300 : Aucune présence détectée sur le canapé
5891	M02	INT (binaire)	livingroom/ ambiance/ motion	De 02- 2020 à 08-2020	1 = Présence détectée 0 = Aucune présence détectée
5892	M03	INT (binaire)	bedroom/ ambiance/ motion	De 02- 2020 à 08-2020	1 = Présence détectée 0 = Aucune présence détectée
5893	M05	INT (binaire)	kitchen/ ambiance/ motion	De 02- 2020 à 08-2020	1 = Présence détectée 0 = Aucune présence détectée

Tableau 3 : Tableau récapitulant les capteurs, leur référence, leur type, les plages de données relevées, ainsi que leurs états

## Tableau des observations graphiques menées pour évaluer la présence dans l'appartement

Ceci fait référence à la partie suivante du rapport : 2.3.3 Analyse graphique des données

N°	Capteurs	Pièces	Attendus	Observations	Evénements
<b>Concordance de l'entrée dans l'appartement avec la présence dans l'entrée (et les sorties)</b>					
1	5888 (D01) 5894 (M01)	Entrance / Corridor	Concordance présence dans l'entrée après entrée dans l'appartement	On trouve une corrélation entre l'ouverture de la porte et la présence dans l'entrée (cf. image 1). Mais certaines observations ne confirment pas cela, plusieurs hypothèses : ouverture répétées successives de la porte sans sortir ; présence de plusieurs personnes dans l'appartement de manière temporaire.	Ev : Entrer dans l'appartement Sortir de l'appartement Ctxt : Habitant présent dans l'appartement Habitant absent de l'appartement <b>Autres éléments : ouvrir la porte à quelqu'un et rester sur le perron sans sortir ; quelqu'un est invité dans l'appartement</b>
<b>Conclusions</b> La corrélation de ses deux capteurs peut indiquer les événements suivants :					



- L'habitant entre dans l'appartement : ouverture de la porte suivie de mouvement dans l'entrée
  - L'habitant sort de l'appartement : ouverture de la porte suivie d'absence de mouvement dans l'entrée
- D'autres événements sont à prendre en compte :
- Quelqu'un est invité dans l'appartement et / ou l'habitant vérifie la fermeture de sa porte d'entrée (en ouvrant et fermant la porte) : ouverture et fermeture de la porte très rapprochées et présence (importante éventuellement si plus d'une personne ?) dans l'entrée
  - L'habitant ouvre la porte et discute sur le perron sans inviter la personne à rentrer : ouverture et fermeture de la porte rapprochées, mouvements **possibles mais non nécessaires** entre ces ouvertures / fermetures
- Critique de l'analyse :
- Les données étant extraites à la minute, certaines valeurs (pics) établissant la corrélation des résultats graphiquement pourraient être absentes et fausser l'analyse à un instant t.
  - En restant sur l'hypothèse principale de la présence d'une unique personne dans l'appartement, ce sont bien les capteurs de mouvement des différentes pièces qui indiquent la présence ou non de l'habitant dans l'appartement (hypothèse vérifiée au point n°2 « Présence globale »).
  - L'analyse de la présence globale pourra nous indiquer si cette analyse de mouvement dans l'entrée et de la porte est suffisante pour tirer les conclusions exposées ci-dessus.

Présence globale					
2	5894 (M01) 5891 (M02) 5892 (M03) 5893 (M05) 5895 (M06)	Corridor Livingroom Bedroom Kitchen Bathroom	Présence globale dans l'appartement. Hypothèse : appartement occupé par un unique habitant	On observe une grande présence dans l'appartement. Il semble que l'habitant ne quitte pas beaucoup l'appartement (2020 : période de confinement covid / télétravail / chômage ?). De plus, on observe une absence de mouvement la nuit (l'habitant dort, cf. image 2). Néanmoins, on peut faire constat de données qui paraissent incohérentes, comme une présence importante détectée la nuit dans la salle de bain (problème du capteur ?)	Ev / Ctxt : Habitant présent dans l'appartement Habitant absent de l'appartement
3	5894 (M01) 5891 (M02) 5892 (M03) 5893 (M05) 5895 (M06) 5888 (D01)	Corridor Livingroom Bedroom Kitchen Bathroom Entrance	Présence globale dans l'appartement. Hypothèse : appartement occupé par un unique habitant Vérifications plus précises des entrées / sorties avec les contacts de la porte	Hypothèse de plusieurs personnes dans l'appartement forte (image 5). Incohérence dans la visualisation des données : porte d'entrée utilisée mais pas de signe de présence dans le couloir, mais plus tard signe de présence dans d'autres pièces (est-ce un manque de données, qualité faible du capteur ?)	Ev : Entrer dans l'appartement Sortir de l'appartement  Ctxt : Habitant présent dans l'appartement Habitant absent de l'appartement

## Conclusions

Il semble difficile de tirer des conclusions malgré le fait de tirer la précision des analyses jusqu'à une remontée des informations à 15 secondes. Voici ce qui est constaté :

- Absence de mouvement la nuit ou quelques mouvements, principalement dans la chambre et dans la salle de bain. On peut en convenir que la personne bouge la nuit quand elle dort ou bien qu'elle se lève pour aller dans la salle de bain. Aucune présence n'est détectée, la plupart du temps, dans ce cas dans le couloir alors que c'est un passage obligé entre chaque pièce.
- Relevé d'incohérences sur certaines activités compte-tenu de la temporalité : par exemple, grande activité dans la salle de bain (des heures), pendant la nuit, sans détection dans le couloir pour le changement de pièce par la suite. On peut suspecter qu'un problème est survenu sur le capteur.
- L'habitant ne semble pas sortir de l'appartement : les détections de la porte d'entrées sont quasi-toujours suivi de mouvements dans d'autres pièces plusieurs minutes après l'action sur la porte, sans détection dans le couloir (incohérence des résultats/manque de données ?). On relève aussi des contacts avec la porte avec détection de l'habitant dans différentes pièces avant et après cette action. Trois hypothèses envisagées :
  - o L'habitant reçoit quelqu'un sur le perron de sa porte mais ne sort pas
  - o L'habitant reçoit quelqu'un chez lui et donc ne sort pas, mais c'est bien l'invité qui entre et qui sort
  - o Il y a plusieurs personnes qui vivent dans l'appartement (contrairement à ce que le rapport sur le dataset indique).

Il y est difficile de déterminer quelle est la situation ici.

Nous sommes allés jusqu'à échantillonner les données toutes les 15 secondes pour l'analyse de la présence (et surtout des entrées/sorties), afin d'être le plus précis possible et de manquer le moins de données possibles. Il est possible que la certaine incohérence relevée, notamment des ouvertures/fermetures de portes peuvent être manquées, ce qui pourrait expliquer certaines situations. Il n'est pas possible de traiter les données avec une fréquence d'échantillonnage plus faible, sinon de le faire sur une quantité de données restreinte.

Les données étant remontées à la seconde pour chaque capteur, il peut s'agir aussi d'un problème dans la duplication de la donnée à la minute  $m+1$  qui ne correspond pas en réalité à l'état à cet instant.

On peut, en dernier lieu, supposer aussi une défaillance technique des capteurs qui n'auraient pas relevé l'information d'une présence à un instant  $t$  particulier durant lequel il y aurait eu une présence, et vis-versa.

### Succession d'entrées / présence dans les pièces

4	5891 (M02) 6220 (D02)	Livingroom Balcon	Passage de différentes pièces : entrée sur le balcon ou sortie	On observe clairement les différents passages entre le balcon et le livingroom, avec des ouvertures et fermetures de portes du balcon avec absence de mouvement dans le livingroom ce qui insinue la présence sur le balcon. L'analyse étant réalisée sur des données remontées toutes les 15s puis toutes les secondes, on peut appuyer une critique sur la remonter des données du capteurs, notamment l'absence de présence de mouvement <b>directement</b>	Ev : Entrer sur le balcon Sortir du balcon Ctxt : Présence sur le balcon / Absence du livingroom Présence dans le livingroom / Absence du balcon
---	--------------------------------	----------------------	--	--	---

				après ce qui est suspecté d'être une fermeture de porte, la remontée des données de mouvements détectés étant présentes mais plus tardives ce qui semble curieux.	
5	5891 (M02) 5894 (M01) 5895 (M06)	Livingroom Corridor Bathroom	Passage de différentes pièces : du living room à la salle de bain en passant par le couloir et vis-versa	Difficile de tirer des observations pertinentes (cf. image 9), on observe des groupements de données mais à des fréquences et horaires différents.	
6	5891 (M02) 5894 (M01) 5893 (M05)	Livingroom Corridor Kitchen	Passage de différentes pièces : du living room à la cuisine en passant par le couloir et vis-versa	Difficile de tirer des observations pertinentes (cf. image 10), on observe des groupements de données, avec des cibles horaires telles que le matin-midi-soir(20h) pour la cuisine, le soir pour le livingroom, mais il y a quand même des activités importantes l'après-midi dans les deux pièces (dans l'hypothèse de quelqu'un qui travaillerait à l'extérieur de chez lui). <b>Il peut s'agir éventuellement du week-end ou d'un jour à la maison, de quelqu'un qui fait du télétravail ou qui se trouve confiné (2020).</b>	
7	5891 (M02) 5894 (M01) 5892 (M03)	Livingroom Corridor Bedroom	Passage de différentes pièces : du living room à la chambre en passant par le couloir et vis-versa	Difficile de tirer des observations pertinentes (cf. image 11), on observe des groupements de données qui reviennent mais souvent à des horaires différents, ou des présences qui paraissent anormales dans la chambre.	
8	5893 (M05) 5894 (M01) 5895 (M06)	Kitchen Corridor Bathroom	Passage de différentes pièces : de la cuisine à la salle de bain en passant par le	Mêmes commentaires que précédemment Cf. image 12 : grande activité dans la cuisine certains midi uniquement	

			couloir et vis-versa		
9	5893 (M05) 5894 (M01) 5892 (M03)	Kitchen Corridor Bedroom	Passage de différentes pièces : de la cuisine à la chambre en passant par le couloir et vis-versa	Mêmes commentaires que précédemment	
10	5892 (M03) 5894 (M01) 5895 (M06)	Bedroom Corridor Bathroom	Passage de différentes pièces : de la chambre à la salle de bain en passant par le couloir et vis-versa	Mêmes commentaires que précédemment	

### Conclusions

Difficile de tirer des conclusions par l'analyse graphique. Une analyse plus quantitative est donc plus pertinente pour tenter de tirer des généralités. Les *a priori* qu'on pourrait avoir sur un comportement « normal » ou « stéréotype » sont à manier avec précaution.

### Concordance d'activités

11	5892 (M03) 5896 (P02) 5895 (M06)	Bedroom (Motion) Bedroom (Bed Pressure) Bathroom	Déterminer la concordance de la présence dans la chambre la nuit avec la pression sur le lit	Concordance pression du lit et présence dans les pièces notamment la chambre et la salle de bain. <b>MAIS</b> activité anormale dans la salle de bain la nuit à partir de la mi-mars	
12	5892 (M03) 6636 (SMP05) 5895 (M06)	Bedroom (Motion) Bathroom (Washing Machine) Bathroom	Déterminer la raison (machine à laver) de la grande activité de la salle de bain sur certaines périodes en confirmant que l'habitant est endormi dans son lit	Pas de corrélations entre la machine à laver et les mouvements importants dans la salle de bain observés à partir de la mi-mars. On en conclut un défaut majeur du capteur ou un problème de duplication de la donnée. Décision : nous ferons un traitement particulier des données majoritairement du début de la collecte (fin-février) jusqu'à l'observation de ce défaut (mi-mars), pour les défauts observés	
13	5891 (M02)	Livingroom Livingroom (TV)	Déterminer l'activité « regarder la	Concordance de l'utilisation de la télé avec le canapé dans le	

6127 (L01) 5889 (D01)	Livingroom (Couch)	télé » en concordance avec la présence dans la pièce et la pression sur le canapé	livingroom le plus souvent l'après-midi, jusqu'au soir	
--------------------------------	-----------------------	--	---	--

Tableau 4 : Tableau des analyses menées à partir de visualisation des données uniquement pour évaluer la présence dans l'appartement

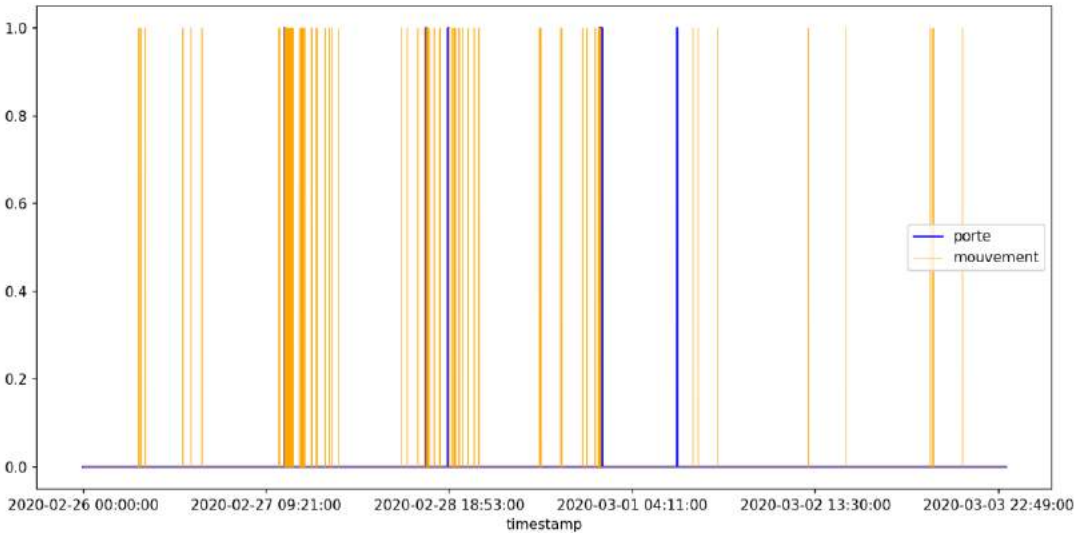


Figure 9 : [image 1] Graphique de la porte d'entrée et mouvement dans l'entrée (données à la minute sur une semaine)

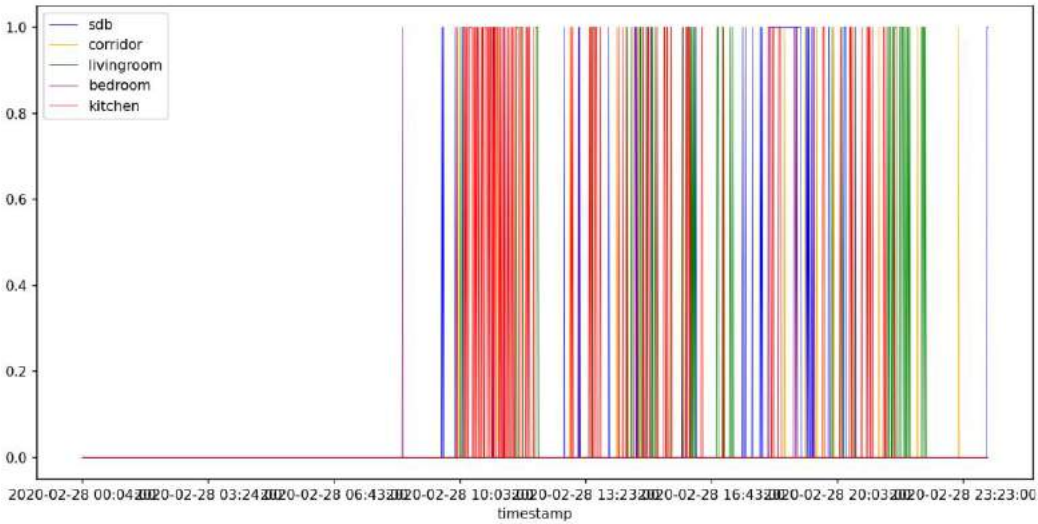


Figure 10 : [image 2] Graphique des capteurs de mouvement (données à la minute sur une journée)

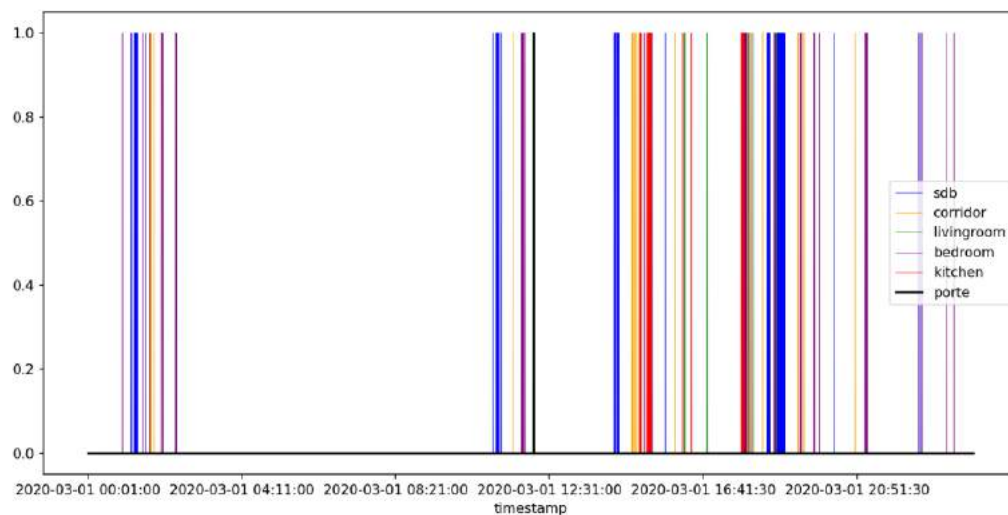


Figure 11 : [image 5] Graphique des capteurs de mouvement et porte d'entrée (données toutes les 15 sec sur la journée)

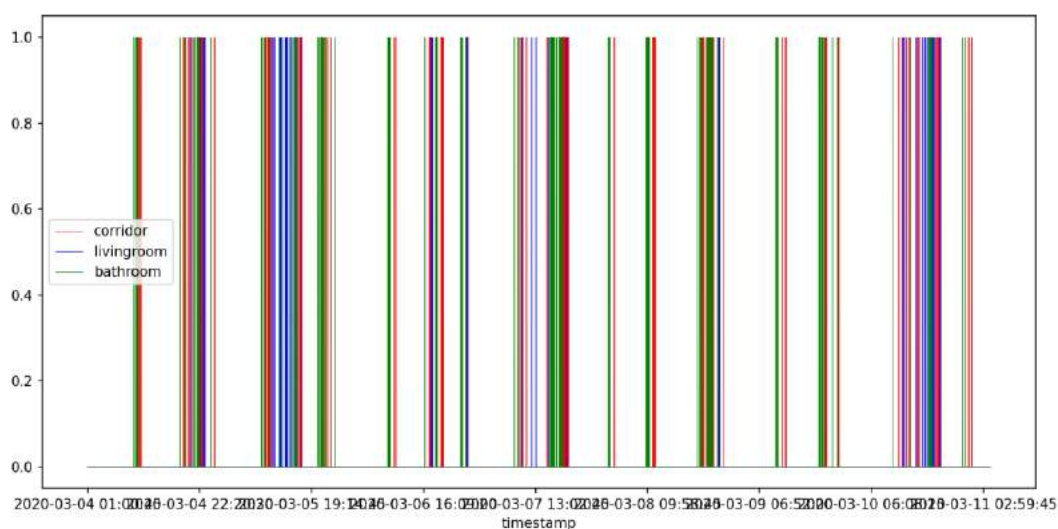


Figure 12 : [image 9] Graphique des capteurs de mouvement livingroom / bathroom / couloir (données toutes les 15 sec sur la semaine)

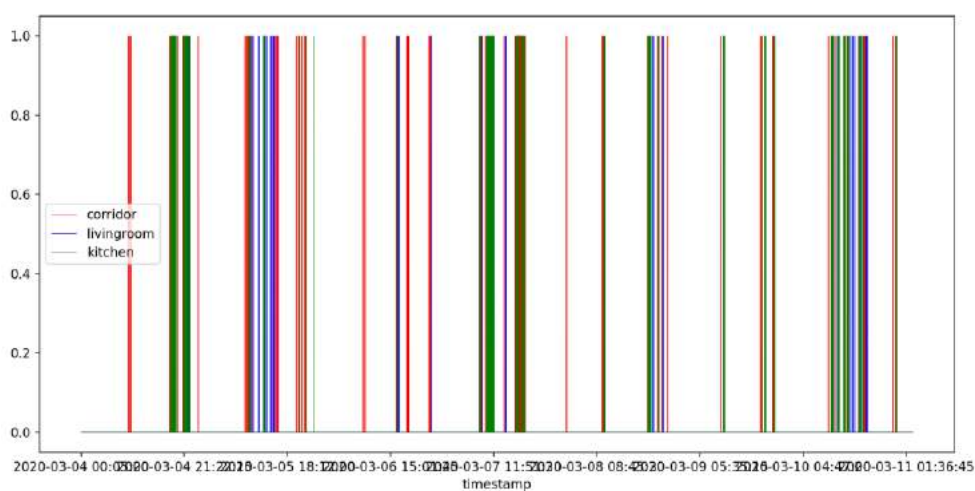


Figure 13 : [image 10] Graphique des capteurs de mouvement livingroom / kitchen / couloir (données toutes les 15 sec sur la semaine)

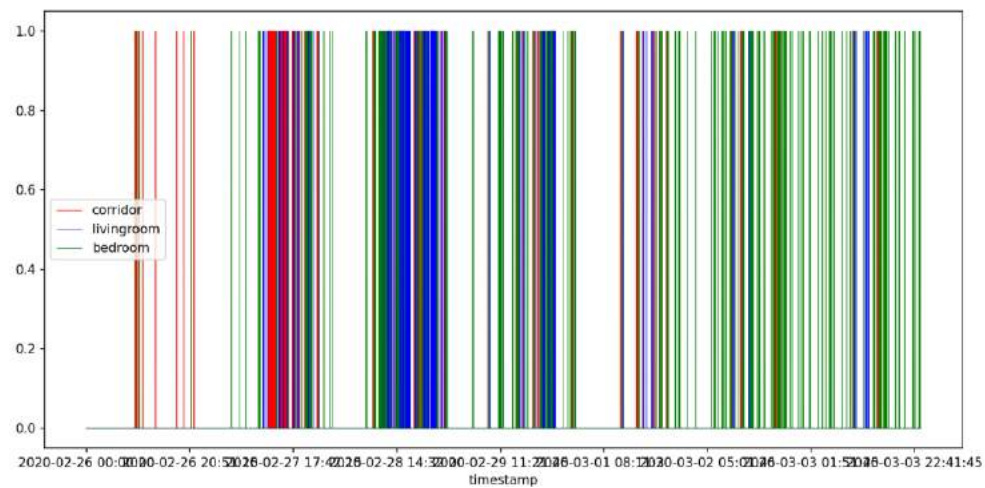


Figure 14 : [image 11] Graphique des capteurs de mouvement livingroom / bedroom / couloir (données toutes les 15 sec sur la semaine)

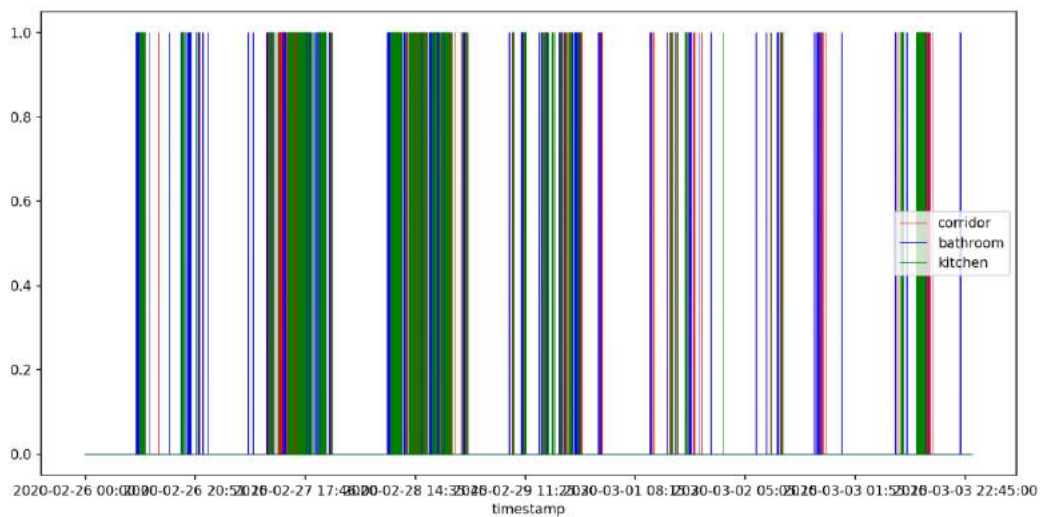


Figure 15 : [image 12] Graphique des capteurs de mouvement bathroom / kitchen / couloir (données toutes les 15 sec sur la semaine)

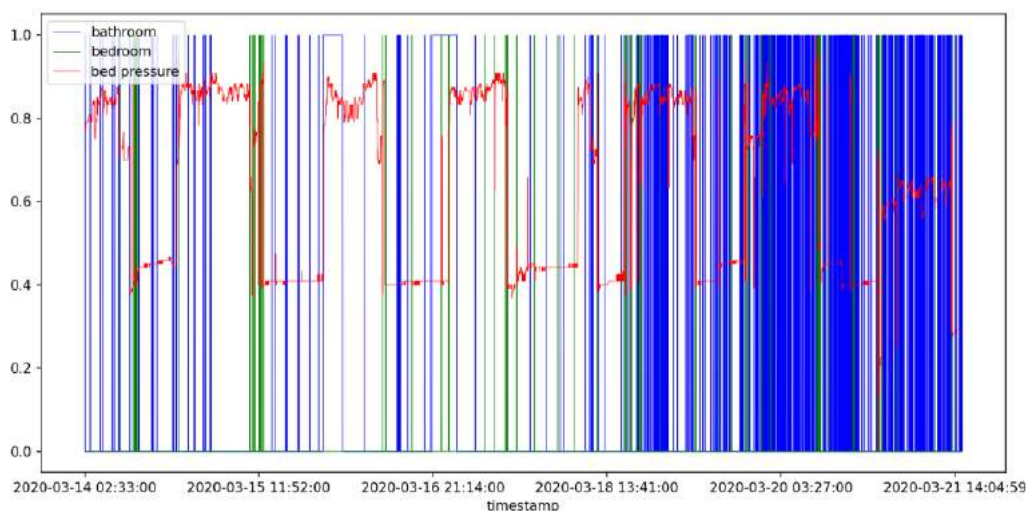


Figure 16 : [image 13] Graphique des capteurs de mouvement bathroom / bedroom et pression du lit (données toutes les minutes sur la semaine)

## Résultats obtenus pour l'analyse de l'activité « regarder la télé »

Ceci fait référence à la partie suivante du rapport : 2.3.4.1 Activité « regarder la télé »



Figure 17 : Visualisation des plages de l'activité "regardé la télé" pour l'ensemble des données

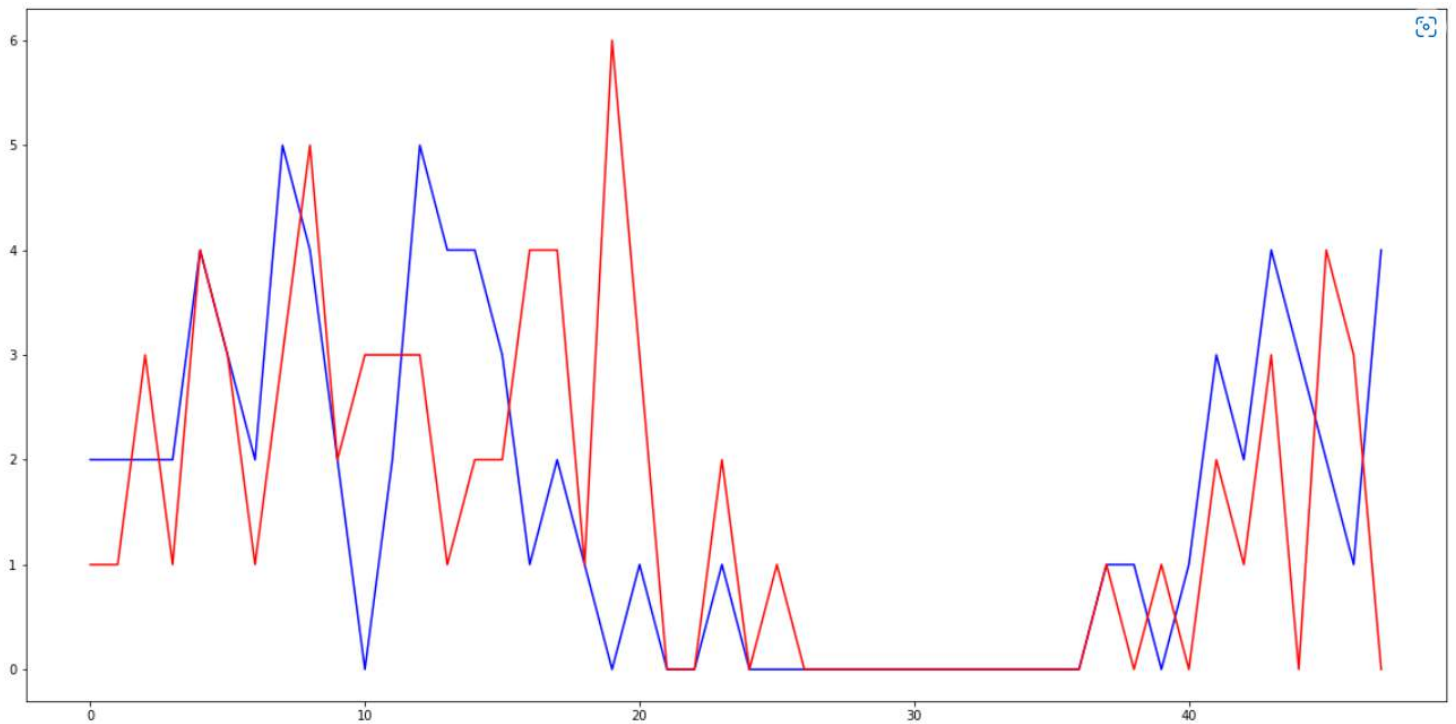


Figure 18 : Visualisation de la fréquence selon le début et la fin de l'activité "regarder la télé"

N.B. : L'indice 0 du graphe correspondant à 15 :00, le graphique couvrant 24 heures



	créneau	durée_moy	fréquence				
0	0:00:00	NaN	0	17	8:30:00	NaN	0
1	0:30:00	NaN	0	18	9:00:00	NaN	0
2	1:00:00	1:24:00	1	19	9:30:00	NaN	0
3	1:30:00	NaN	0	20	10:00:00	0:24:00	1
4	2:00:00	NaN	0	21	10:30:00	NaN	0
5	2:30:00	0:47:00	1	22	11:00:00	0:34:00	1
6	3:00:00	NaN	0	23	11:30:00	2:39:30	2
7	3:30:00	NaN	0	24	12:00:00	0:22:00	1
8	4:00:00	NaN	0	25	12:30:00	1:34:30	2
9	4:30:00	NaN	0	26	13:00:00	3:28:00	1
10	5:00:00	NaN	0	27	13:30:00	NaN	0
11	5:30:00	NaN	0	28	14:00:00	4:35:00	2
12	6:00:00	NaN	0	29	14:30:00	0:58:40	3
13	6:30:00	NaN	0	30	15:00:00	3:23:00	1
14	7:00:00	NaN	0	31	15:30:00	4:28:01	1
15	7:30:00	NaN	0	32	16:00:00	5:11:00	2
16	8:00:00	NaN	0	33	16:30:00	0:34:30	2
				34	17:00:00	2:33:00	3
				35	17:30:00	4:00:00	2
				36	18:00:00	2:42:00	1
				37	18:30:00	1:48:00	4
				38	19:00:00	1:42:00	4
				39	19:30:00	0:44:00	2
				40	20:00:00	NaN	0
				41	20:30:00	NaN	0
				42	21:00:00	2:45:40	3
				43	21:30:00	1:03:20	3
				44	22:00:00	1:56:00	4
				45	22:30:00	1:47:00	3
				46	23:00:00	NaN	0
				47	23:30:00	NaN	0

Tableau 5 : Table des fréquences et durées moyennes pour l'activité "regardé la télé"

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim
0:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
0:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
1:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	1:24:00 - 1
1:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
2:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
2:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:47:00 - 1	Nan - 0
3:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
3:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
4:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
4:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
5:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
5:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
6:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
6:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
7:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
7:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
8:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
8:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
9:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
9:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
10:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:24:00 - 1
10:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
11:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:34:00 - 1	Nan - 0	Nan - 0
11:30:00	3:29:01 - 1	Nan - 0	Nan - 0	Nan - 0	1:50:00 - 1	Nan - 0	Nan - 0
12:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:22:00 - 1	Nan - 0	Nan - 0
12:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	1:28:01 - 1	1:41:00 - 1
13:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:28:00 - 1	Nan - 0	Nan - 0
13:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
14:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	8:39:01 - 1	0:31:00 - 1
14:30:00	Nan - 0	Nan - 0	Nan - 0	1:40:00 - 1	Nan - 0	Nan - 0	0:38:00 - 2

15:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:23:00 - 1	Nan - 0	Nan - 0
15:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	4:28:01 - 1	Nan - 0
16:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	5:11:00 - 2	Nan - 0	Nan - 0
16:30:00	0:47:00 - 1	Nan - 0	Nan - 0	0:22:00 - 1	Nan - 0	Nan - 0	Nan - 0
17:00:00	Nan - 0	Nan - 0	1:53:00 - 1	2:53:00 - 2	Nan - 0	Nan - 0	Nan - 0
17:30:00	7:22:00 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	0:38:00 - 1
18:00:00	Nan - 0	Nan - 0	Nan - 0	2:42:00 - 1	Nan - 0	Nan - 0	Nan - 0
18:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	4:28:00 - 1	1:09:00 - 2	0:26:00 - 1
19:00:00	Nan - 0	Nan - 0	5:25:00 - 1	0:25:00 - 1	0:30:00 - 1	Nan - 0	0:28:00 - 1
19:30:00	Nan - 0	Nan - 0	Nan - 0	0:54:00 - 1	0:34:01 - 1	Nan - 0	Nan - 0
20:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
20:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
21:00:00	3:57:01 - 1	Nan - 0	Nan - 0	Nan - 0	0:40:00 - 1	3:40:00 - 1	Nan - 0
21:30:00	Nan - 0	Nan - 0	0:35:00 - 1	1:45:00 - 1	Nan - 0	Nan - 0	0:50:00 - 1
22:00:00	2:40:00 - 1	Nan - 0	2:27:00 - 1	Nan - 0	Nan - 0	0:56:00 - 1	1:41:01 - 1
22:30:00	Nan - 0	2:39:00 - 1	Nan - 0	0:59:00 - 1	1:43:00 - 1	Nan - 0	Nan - 0
23:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
23:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0

Tableau 6 : Calendrier des fréquences et durées moyennes pour l'activité "regardé la télé"

## Résultats obtenus pour l'analyse de l'activité « dormir dans le lit »

Ceci fait référence à la partie suivante du rapport : 2.3.4.2 Activité « dormir dans le lit »

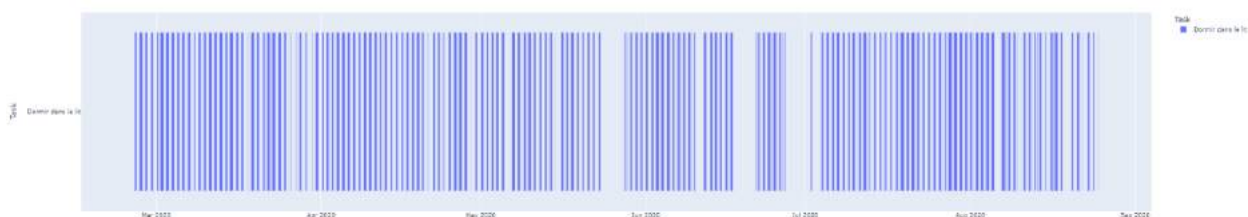


Figure 19 : Visualisation des plages pour l'activité "dormir dans le lit"

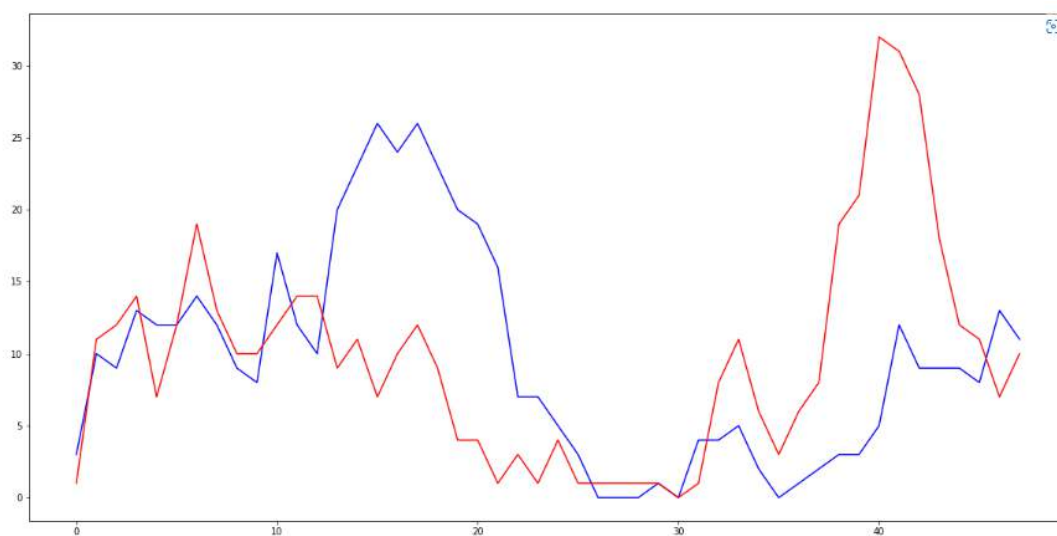


Figure 20 : Visualisation de la fréquence d'appartenance de l'activité à un créneau selon le début de la fin de l'activité "dormir dans le lit", l'indice 0 du graphe correspondant à 15:00

	créneau	durée_moy	fréquence				
0	0:00:00	NaN	0				
1	0:30:00	10:05:21	16				
2	1:00:00	10:16:45	17	25	12:30:00	3:47:00	1
3	1:30:00	9:30:41	15	26	13:00:00	NaN	0
4	2:00:00	9:21:00	4	27	13:30:00	4:07:59	1
5	2:30:00	8:49:25	5	28	14:00:00	3:19:01	1
6	3:00:00	8:32:48	5	29	14:30:00	3:18:01	1
7	3:30:00	8:01:39	3	30	15:00:00	4:11:00	2
8	4:00:00	6:52:57	1	31	15:30:00	3:28:00	1
9	4:30:00	NaN	0	32	16:00:00	NaN	0
10	5:00:00	NaN	0	33	16:30:00	NaN	0
11	5:30:00	NaN	0	34	17:00:00	NaN	0
12	6:00:00	NaN	0	35	17:30:00	NaN	0
13	6:30:00	NaN	0	36	18:00:00	3:09:00	1
14	7:00:00	NaN	0	37	18:30:00	NaN	0
15	7:30:00	4:07:26	2	38	19:00:00	NaN	0
16	8:00:00	5:20:59	1	39	19:30:00	14:42:00	2
17	8:30:00	NaN	0	40	20:00:00	14:48:20	3
18	9:00:00	NaN	0	41	20:30:00	8:48:48	3
19	9:30:00	3:17:00	1	42	21:00:00	15:08:41	3
20	10:00:00	NaN	0	43	21:30:00	13:28:43	11
21	10:30:00	3:29:01	1	44	22:00:00	11:12:24	10
22	11:00:00	NaN	0	45	22:30:00	11:01:28	15
23	11:30:00	NaN	0	46	23:00:00	11:08:37	17
24	12:00:00	3:50:00	1	47	23:30:00	10:07:31	17

Tableau 7 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim
0:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
0:30:00	11:39:00 - 1	10:37:36 - 5	8:17:01 - 1	8:31:53 - 3	10:22:20 - 3	10:30:01 - 1	10:34:30 - 2
1:00:00	11:39:01 - 1	10:31:00 - 2	10:44:00 - 1	10:25:30 - 2	10:35:35 - 2	9:33:09 - 5	10:23:00 - 4
1:30:00	Nan - 0	Nan - 0	9:10:41 - 3	10:09:29 - 2	10:08:00 - 2	9:11:36 - 7	10:12:01 - 1
2:00:00	Nan - 0	9:23:01 - 1	8:47:00 - 1	Nan - 0	9:50:01 - 1	Nan - 0	9:24:00 - 1
2:30:00	Nan - 0	Nan - 0	10:08:02 - 1	6:49:01 - 1	Nan - 0	Nan - 0	9:03:21 - 3
3:00:00	7:41:01 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0	9:48:40 - 3	5:37:00 - 1
3:30:00	6:52:57 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0	8:36:00 - 2	Nan - 0
4:00:00	6:52:57 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
4:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
5:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
5:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
6:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
6:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
7:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
7:30:00	3:58:54 - 1	Nan - 0	Nan - 0	4:15:59 - 1	Nan - 0	Nan - 0	Nan - 0
8:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	5:20:59 - 1	Nan - 0
8:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
9:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
9:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:17:00 - 1
10:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
10:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:29:01 - 1	Nan - 0
11:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
11:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
12:00:00	Nan - 0	Nan - 0	3:50:00 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0
12:30:00	Nan - 0	Nan - 0	Nan - 0	3:47:00 - 1	Nan - 0	Nan - 0	Nan - 0
13:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
13:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	4:07:59 - 1	Nan - 0	Nan - 0
14:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:19:01 - 1	Nan - 0	Nan - 0
14:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:18:01 - 1
15:00:00	Nan - 0	Nan - 0	Nan - 0	4:57:00 - 1	Nan - 0	Nan - 0	3:25:00 - 1
15:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:28:00 - 1	Nan - 0	Nan - 0
16:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
16:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
17:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
17:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
18:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	3:09:00 - 1	Nan - 0
18:30:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
19:00:00	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0	Nan - 0
19:30:00	Nan - 0	Nan - 0	15:14:01 - 1	Nan - 0	Nan - 0	Nan - 0	14:10:00 - 1
20:00:00	14:21:59 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0	14:28:00 - 1	15:35:02 - 1
20:30:00	Nan - 0	Nan - 0	Nan - 0	11:35:42 - 2	Nan - 0	Nan - 0	3:15:01 - 1
21:00:00	14:15:00 - 1	16:30:03 - 1	Nan - 0	Nan - 0	Nan - 0	Nan - 0	14:41:00 - 1
21:30:00	13:48:59 - 2	13:02:00 - 1	13:10:30 - 2	13:21:00 - 2	14:52:00 - 1	14:54:01 - 1	12:23:31 - 2
22:00:00	11:27:30 - 2	13:33:01 - 1	13:22:01 - 1	11:13:00 - 2	Nan - 0	7:31:30 - 2	12:22:29 - 2
22:30:00	12:23:00 - 2	11:12:30 - 2	11:05:40 - 3	12:03:30 - 2	10:35:30 - 2	8:15:29 - 2	11:32:30 - 2
23:00:00	12:21:00 - 2	11:57:02 - 4	7:49:00 - 1	12:05:58 - 2	9:53:31 - 2	12:41:30 - 2	9:56:21 - 4
23:30:00	11:13:30 - 4	10:22:00 - 1	10:08:00 - 2	9:55:54 - 3	11:30:00 - 1	12:31:41 - 3	5:54:21 - 3

Tableau 8 : Calendrier des moyennes de durées et des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de plage

Résultats obtenus pour l'analyse de la corrélation entre deux activités :  
« regarder la télé » et « dormir dans le lit »

Ceci fait référence à la partie suivante du rapport : 2.3.4.3 *Corrélation des activités « regarder la télé » et « dormir dans le lit »*

	début_rgtv	fin_rgtv	durée_rgtv	début_sleep	fin_sleep	durée_sleep
0	2020-02-27 18:47:00	2020-02-27 19:12:00	0:25:00	2020-02-27 19:15:00	2020-02-27 19:22:01	0:07:01
1	2020-02-27 21:38:00	2020-02-27 22:49:00	1:11:00	2020-02-27 22:53:00	2020-02-28 10:04:57	11:11:57
2	2020-03-29 21:52:00	2020-03-29 23:33:01	1:41:01	2020-03-29 23:36:00	2020-03-29 23:38:00	0:02:00
3	2020-04-01 22:08:00	2020-04-02 00:35:00	2:27:00	2020-04-02 00:44:00	2020-04-02 09:47:00	9:03:00
4	2020-04-02 22:33:00	2020-04-02 23:32:00	0:59:00	2020-04-02 23:37:00	2020-04-03 10:45:00	11:08:00
5	2020-04-03 16:13:00	2020-04-03 19:11:00	2:58:00	2020-04-03 19:12:00	2020-04-03 19:25:01	0:13:01
6	2020-04-03 20:46:00	2020-04-03 21:26:00	0:40:00	2020-04-03 21:27:00	2020-04-04 12:19:00	14:52:00
7	2020-04-04 20:59:00	2020-04-04 21:08:00	0:09:00	2020-04-04 21:23:00	2020-04-05 12:17:01	14:54:01
8	2020-04-07 22:27:00	2020-04-08 01:06:00	2:39:00	2020-04-08 01:19:00	2020-04-08 10:52:58	9:33:58
9	2020-04-26 20:52:00	2020-04-26 20:55:00	0:03:00	2020-04-26 20:59:00	2020-04-27 11:40:00	14:41:00
10	2020-05-21 21:23:00	2020-05-21 23:08:00	1:45:00	2020-05-21 23:24:00	2020-05-21 23:32:00	0:08:00
11	2020-05-21 23:20:00	2020-05-21 23:22:00	0:02:00	2020-05-21 23:24:00	2020-05-21 23:32:00	0:08:00
12	2020-05-22 10:59:00	2020-05-22 11:33:00	0:34:00	2020-05-22 11:37:53	2020-05-22 12:31:00	0:53:07
13	2020-05-23 02:38:00	2020-05-23 03:25:00	0:47:00	2020-05-23 03:30:00	2020-05-23 12:10:00	8:40:00
14	2020-05-24 00:57:00	2020-05-24 02:21:00	1:24:00	2020-05-24 02:25:00	2020-05-24 02:27:59	0:02:59
15	2020-06-01 21:06:59	2020-06-02 01:04:00	3:57:01	2020-06-02 01:07:00	2020-06-02 12:41:00	11:34:00
16	2020-06-03 17:05:00	2020-06-03 18:58:00	1:53:00	2020-06-03 19:03:00	2020-06-03 19:15:00	0:12:00
17	2020-06-22 17:21:00	2020-06-23 00:43:00	7:22:00	2020-06-23 00:49:00	2020-06-23 10:17:01	9:28:01
18	2020-07-01 19:06:00	2020-07-02 00:31:00	5:25:00	2020-07-02 00:41:58	2020-07-02 07:32:39	6:50:41
19	2020-07-04 13:59:00	2020-07-04 22:38:01	8:39:01	2020-07-04 22:42:00	2020-07-04 22:49:00	0:07:00
20	2020-07-05 14:15:00	2020-07-05 14:46:00	0:31:00	2020-07-05 14:48:00	2020-07-05 15:03:00	0:15:00
21	2020-07-05 15:07:00	2020-07-05 16:12:00	1:05:00	2020-07-05 16:19:00	2020-07-05 16:50:00	0:31:00

Tableau 9 : Table des plages des deux activités "regarder la télé" et "dormir dans le lit" qui se suivent à 20 minutes

## Résultats obtenus pour l'analyse de l'activité « faire un repas »

Ceci fait référence à la partie suivante du rapport : 2.3.4.4 Activité « faire un repas »

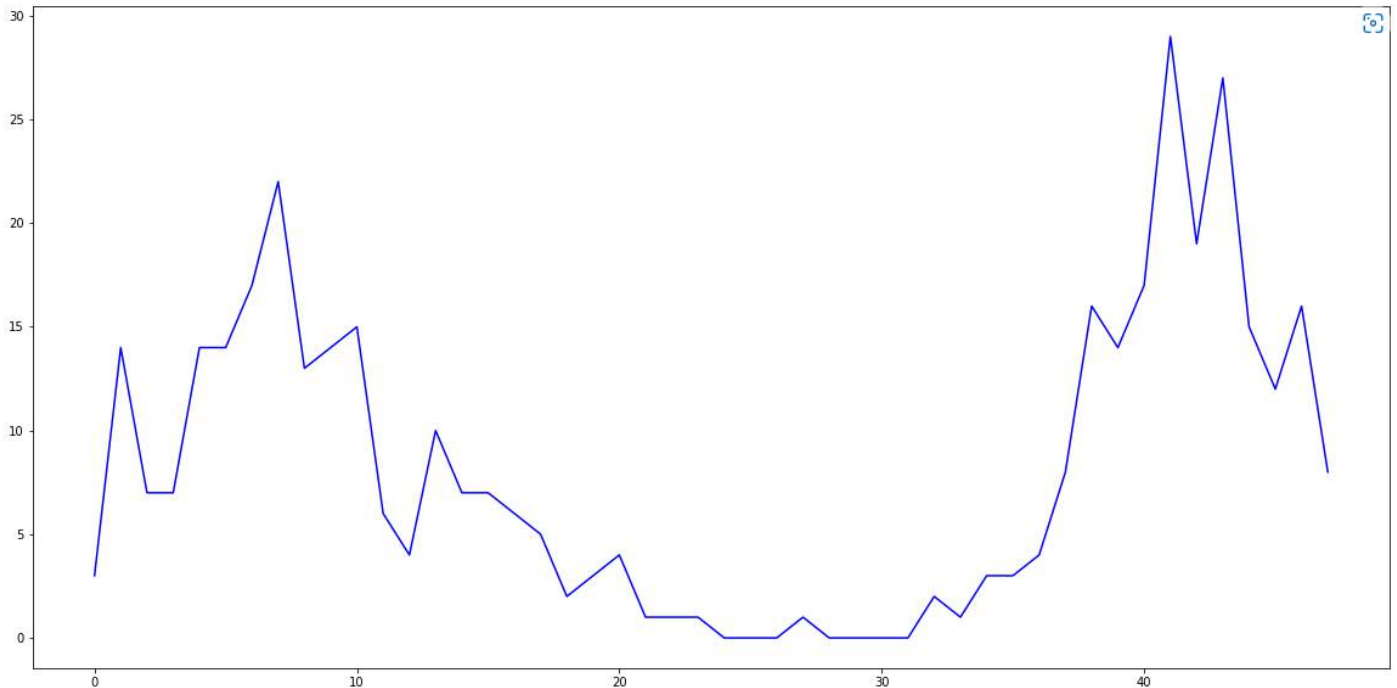


Figure 21 : Visualisation des fréquences de l'activité "faire un repas" par la combinaison de 3 capteurs (frigo, machine à café, sandwichmaker), en considérant les données du frigo de 1 sec

Une deuxième analyse a été réalisée n'incluant pas les données du frigo à 1 seconde :

[1. Pourcentages par rapport aux nombres de plages : 97]

{2. Pourcentages par rapport aux nombres de créneaux totaux : 48}

- Les pics de fréquence les plus importants sont observés aux horaires suivants à 14 :00 (9) [soit 9%], puis 12 :00 (8) [soit 8,2%], 11 :00-11 :30 (6) [soit 12,4%], ainsi que 10 :30 / 12 :30 / 20 :00 (5) [soit 15,5%], soit un total de 45,3% de l'activité.
- On observe encore une certaine activité à 13 :00 / 17 :30 / 18 :30 / 21 :30 [soit 16,5%] c'est-à-dire un total de 61,9% sur toute l'activité relevée.

Ces premiers relevés que l'on peut facilement associés à des périodes de repas dans la journée (le repas du midi et le repas du soir), sont alors cohérents pour cette activité. De plus, on note les observations suivantes :

- Absence d'activité certaine (fréquence de 0) de 01 :00 à 6 :30 {soit 23% de créneaux vides}.
- Absence d'activité étendue (fréquence de 0 ou 1) de 00 :00 à 08 :30 {soit 36% de créneaux vides}. Ainsi, il n'y a pas d'activité de cuisine avec ces appareils la nuit, ce qui est cohérent avec l'activité.



### Machine à café :

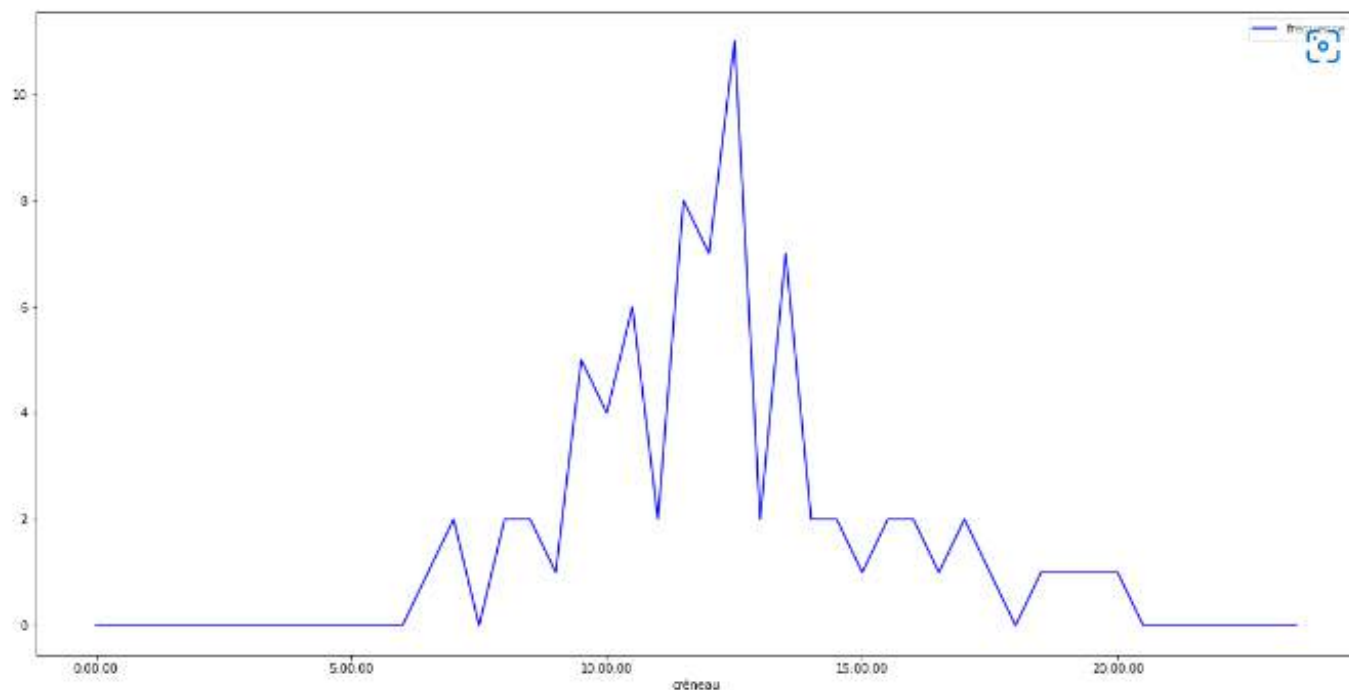


Figure 22 : Visualisation de la fréquence de l'activité de la machine à café correspondant au début de l'activité

	créneau	durée_moy	fréquence	15	7:30:00	NaN	0	32	16:00:00	0:00:30	2
0	0:00:00	NaN	0	16	8:00:00	0:00:30	2	33	16:30:00	0:00:01	1
1	0:30:00	NaN	0	17	8:30:00	0:00:01	2	34	17:00:00	0:00:01	2
2	1:00:00	NaN	0	18	9:00:00	0:00:01	1	35	17:30:00	0:00:01	1
3	1:30:00	NaN	0	19	9:30:00	0:00:01	5	36	18:00:00	NaN	0
4	2:00:00	NaN	0	20	10:00:00	0:00:01	4	37	18:30:00	0:00:01	1
5	2:30:00	NaN	0	21	10:30:00	0:10:10	6	38	19:00:00	0:00:01	1
6	3:00:00	NaN	0	22	11:00:00	0:00:01	2	39	19:30:00	0:00:01	1
7	3:30:00	NaN	0	23	11:30:00	0:00:25	8	40	20:00:00	0:01:00	1
8	4:00:00	NaN	0	24	12:00:00	0:00:09	7	41	20:30:00	NaN	0
9	4:30:00	NaN	0	25	12:30:00	0:00:39	11	42	21:00:00	NaN	0
10	5:00:00	NaN	0	26	13:00:00	0:00:01	2	43	21:30:00	NaN	0
11	5:30:00	NaN	0	27	13:30:00	0:00:09	7	44	22:00:00	NaN	0
12	6:00:00	NaN	0	28	14:00:00	0:00:01	2	45	22:30:00	NaN	0
13	6:30:00	0:00:01	1	29	14:30:00	0:00:30	2	46	23:00:00	NaN	0
14	7:00:00	0:00:01	2	30	15:00:00	0:00:01	1	47	23:30:00	NaN	0
15	7:30:00	NaN	0	31	15:30:00	0:00:01	2				

Tableau 10 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim									
								8:30:00	0	1	0	0	0	1	0	
0:00:00	0	0	0	0	0	0	0	9:00:00	0	0	0	1	0	0	0	
0:30:00	0	0	0	0	0	0	0	9:30:00	3	0	0	1	0	0	1	
1:00:00	0	0	0	0	0	0	0	10:00:00	1	0	0	3	0	0	0	
1:30:00	0	0	0	0	0	0	0	10:30:00	0	1	2	1	0	2	0	
2:00:00	0	0	0	0	0	0	0	11:00:00	2	0	0	0	0	0	0	
2:30:00	0	0	0	0	0	0	0	11:30:00	1	2	0	0	2	1	2	
3:00:00	0	0	0	0	0	0	0	12:00:00	0	1	2	0	1	1	2	
3:30:00	0	0	0	0	0	0	0	12:30:00	0	2	0	0	4	4	1	
4:00:00	0	0	0	0	0	0	0	13:00:00	0	0	0	0	1	1	0	
4:30:00	0	0	0	0	0	0	0	13:30:00	2	1	1	1	1	0	1	
5:00:00	0	0	0	0	0	0	0	14:00:00	1	0	0	1	0	0	0	
5:30:00	0	0	0	0	0	0	0	14:30:00	1	0	1	0	0	0	0	
6:00:00	0	0	0	0	0	0	0	15:00:00	1	0	0	0	0	0	0	
6:30:00	0	0	0	0	0	1	0	15:30:00	0	1	0	1	0	0	0	
7:00:00	1	0	0	0	1	0	0	16:00:00	0	0	0	2	0	0	0	
7:30:00	0	0	0	0	0	0	0	16:30:00	0	0	1	0	0	0	0	
8:00:00	1	0	0	1	0	0	0	17:00:00	0	0	0	1	0	1	0	
								17:30:00	0	0	0	0	0	0	0	1
								18:00:00	0	0	0	0	0	0	0	0
								18:30:00	0	0	0	0	1	0	0	0
								19:00:00	1	0	0	0	0	0	0	0
								19:30:00	1	0	0	0	0	0	0	0
								20:00:00	0	0	0	1	0	0	0	0
								20:30:00	0	0	0	0	0	0	0	0
								21:00:00	0	0	0	0	0	0	0	0
								21:30:00	0	0	0	0	0	0	0	0
								22:00:00	0	0	0	0	0	0	0	0
								22:30:00	0	0	0	0	0	0	0	0
								23:00:00	0	0	0	0	0	0	0	0
								23:30:00	0	0	0	0	0	0	0	0

Tableau 11 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de plage



## Lave-vaisselle :

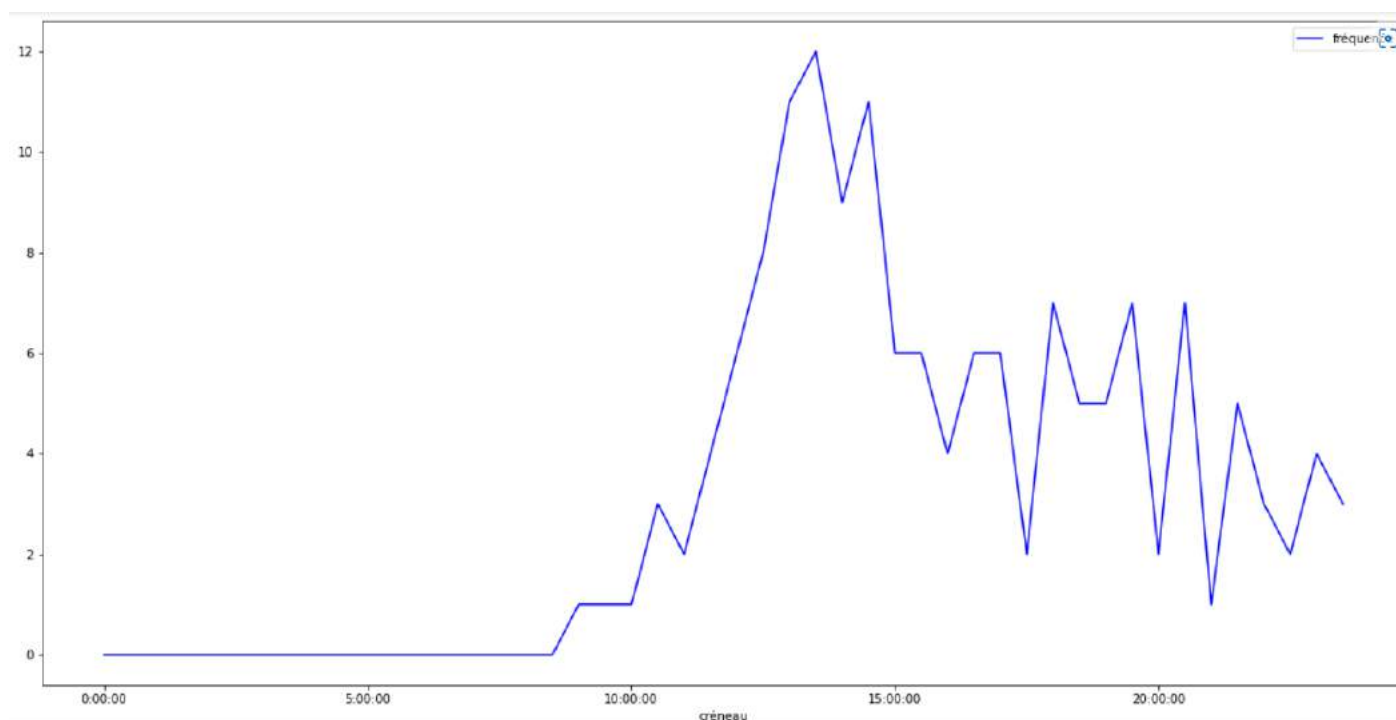


Figure 23 : Visualisation de la fréquence de l'activité du lave-vaisselle correspondant au début de l'activité

	créneau	durée_moy	fréquence								
				16	8:00:00	NaN	0				
0	0:00:00	NaN	0	17	8:30:00	NaN	0	32	16:00:00	0:03:45	4
1	0:30:00	NaN	0	18	9:00:00	0:14:02	1	33	16:30:00	0:06:40	6
2	1:00:00	NaN	0	19	9:30:00	0:03:00	1	34	17:00:00	0:05:30	6
3	1:30:00	NaN	0	20	10:00:00	0:14:01	1	35	17:30:00	0:09:00	2
4	2:00:00	NaN	0	21	10:30:00	0:07:00	3	36	18:00:00	0:08:01	7
5	2:30:00	NaN	0	22	11:00:00	0:02:28	2	37	18:30:00	0:03:23	5
6	3:00:00	NaN	0	23	11:30:00	0:08:58	4	38	19:00:00	0:07:00	5
7	3:30:00	NaN	0	24	12:00:00	0:08:39	6	39	19:30:00	0:03:26	7
8	4:00:00	NaN	0	25	12:30:00	0:08:38	8	40	20:00:00	0:14:02	2
9	4:30:00	NaN	0	26	13:00:00	0:07:00	11	41	20:30:00	0:06:26	7
10	5:00:00	NaN	0	27	13:30:00	0:04:20	12	42	21:00:00	0:04:00	1
11	5:30:00	NaN	0	28	14:00:00	0:04:33	9	43	21:30:00	0:07:58	5
12	6:00:00	NaN	0	29	14:30:00	0:05:40	11	44	22:00:00	0:03:54	3
13	6:30:00	NaN	0	30	15:00:00	0:08:03	6	45	22:30:00	0:08:00	2
14	7:00:00	NaN	0	31	15:30:00	0:04:40	6	46	23:00:00	0:02:56	4
15	7:30:00	NaN	0	32	16:00:00	0:03:45	4	47	23:30:00	0:03:57	3

Tableau 12 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim										
								<b>8:30:00</b>	0	0	0	0	0	0	0	0	0
<b>0:00:00</b>	0	0	0	0	0	0	0	<b>9:00:00</b>	0	0	0	1	0	0	0	0	0
<b>0:30:00</b>	0	0	0	0	0	0	0	<b>9:30:00</b>	0	0	0	1	0	0	0	0	0
<b>1:00:00</b>	0	0	0	0	0	0	0	<b>10:00:00</b>	0	0	0	0	0	1	0	0	0
<b>1:30:00</b>	0	0	0	0	0	0	0	<b>10:30:00</b>	1	0	0	1	0	1	0	0	0
<b>2:00:00</b>	0	0	0	0	0	0	0	<b>11:00:00</b>	1	0	0	0	0	1	0	0	0
<b>2:30:00</b>	0	0	0	0	0	0	0	<b>11:30:00</b>	0	0	0	0	0	3	1	0	0
<b>3:00:00</b>	0	0	0	0	0	0	0	<b>12:00:00</b>	1	1	0	0	1	1	2	0	0
<b>3:30:00</b>	0	0	0	0	0	0	0	<b>12:30:00</b>	0	3	0	1	2	1	1	0	0
<b>4:00:00</b>	0	0	0	0	0	0	0	<b>13:00:00</b>	1	3	1	2	1	1	2	0	0
<b>4:30:00</b>	0	0	0	0	0	0	0	<b>13:30:00</b>	1	2	1	3	2	1	2	0	0
<b>5:00:00</b>	0	0	0	0	0	0	0	<b>14:00:00</b>	1	2	2	2	1	1	0	0	0
<b>5:30:00</b>	0	0	0	0	0	0	0	<b>14:30:00</b>	2	2	2	1	1	2	1	0	0
<b>6:00:00</b>	0	0	0	0	0	0	0	<b>15:00:00</b>	0	1	1	0	1	3	0	0	0
<b>6:30:00</b>	0	0	0	0	0	0	0	<b>15:30:00</b>	1	0	2	0	0	3	0	0	0
<b>7:00:00</b>	0	0	0	0	0	0	0	<b>16:00:00</b>	0	1	1	0	1	1	0	0	0
<b>7:30:00</b>	0	0	0	0	0	0	0	<b>16:30:00</b>	0	1	1	0	2	2	0	0	0
<b>8:00:00</b>	0	0	0	0	0	0	0	<b>17:00:00</b>	0	1	1	0	2	1	1	0	0
								<b>17:30:00</b>	0	0	1	0	0	0	0	1	0
								<b>18:00:00</b>	1	1	2	0	3	0	0	0	0
								<b>18:30:00</b>	1	1	1	0	1	0	1	0	0
								<b>19:00:00</b>	1	1	1	0	2	0	0	0	0
								<b>19:30:00</b>	2	1	1	0	3	0	0	0	0
								<b>20:00:00</b>	1	1	0	0	0	0	0	0	0
								<b>20:30:00</b>	2	1	0	1	2	0	1	0	0
								<b>21:00:00</b>	0	0	0	1	0	0	0	0	0
								<b>21:30:00</b>	1	1	0	0	0	1	2	0	0
								<b>22:00:00</b>	0	0	0	1	0	1	1	0	0
								<b>22:30:00</b>	1	1	0	0	0	0	0	0	0
								<b>23:00:00</b>	0	1	0	0	0	1	2	0	0
								<b>23:30:00</b>	1	1	0	0	0	0	0	1	0

Tableau 13 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de plage

### Lumière de la cuisinière :

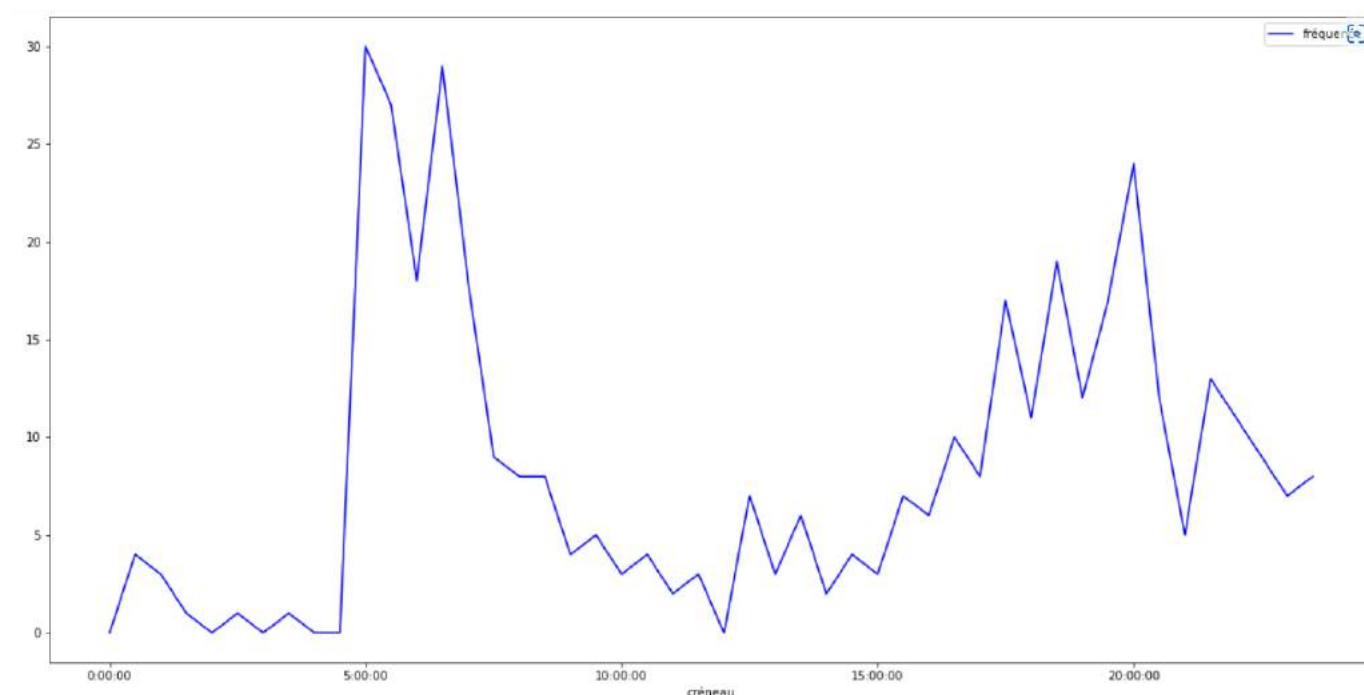


Figure 24 : Visualisation de la fréquence de l'activité correspondant au début de l'activité

	créneau	durée_moy	fréquence	16	8:00:00	6:25:55	8						
0	0:00:00	NaN	0	17	8:30:00	5:30:22	8	33	16:30:00	0:47:22	10		
1	0:30:00	0:35:45	4	18	9:00:00	3:54:00	4		34	17:00:00	0:40:07	8	
2	1:00:00	0:08:40	3	19	9:30:00	5:36:00	5			35	17:30:00	0:24:39	17
3	1:30:00	0:02:00	1	20	10:00:00	4:29:00	3				36	18:00:00	0:16:16
4	2:00:00	NaN	0	21	10:30:00	2:26:59	4	37				18:30:00	0:43:21
5	2:30:00	0:19:00	1	22	11:00:00	0:39:12	2		38			19:00:00	0:28:15
6	3:00:00	NaN	0	23	11:30:00	0:34:20	3			39		19:30:00	0:18:11
7	3:30:00	0:04:01	1	24	12:00:00	NaN	0				40	20:00:00	0:44:50
8	4:00:00	NaN	0	25	12:30:00	2:24:34	7	41				20:30:00	1:11:46
9	4:30:00	NaN	0	26	13:00:00	1:39:00	3		42			21:00:00	0:23:48
10	5:00:00	11:32:03	30	27	13:30:00	4:30:39	6			43		21:30:00	0:20:17
11	5:30:00	11:48:44	27	28	14:00:00	3:40:00	2				44	22:00:00	0:17:27
12	6:00:00	10:15:38	18	29	14:30:00	2:28:42	4	45				22:30:00	0:05:00
13	6:30:00	9:05:49	29	30	15:00:00	1:26:43	3		46			23:00:00	0:08:30
14	7:00:00	7:33:51	18	31	15:30:00	2:48:12	7			47		23:30:00	0:33:07
15	7:30:00	2:42:27	9	32	16:00:00	1:01:50	6						

Tableau 14 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim								
								9:00:00	0	0	0	0	0	1	3
0:00:00	0	0	0	0	0	0	0	9:30:00	0	1	0	0	0	2	2
0:30:00	0	0	2	1	0	1	0	10:00:00	0	0	0	0	1	1	1
1:00:00	0	1	0	0	0	2	0	10:30:00	0	0	2	0	1	0	1
1:30:00	0	0	0	0	1	0	0	11:00:00	0	1	0	0	0	0	1
2:00:00	0	0	0	0	0	0	0	11:30:00	0	0	1	0	1	0	1
2:30:00	0	0	0	0	0	1	0	12:00:00	0	0	0	0	0	0	0
3:00:00	0	0	0	0	0	0	0	12:30:00	0	1	4	0	1	0	1
3:30:00	0	0	0	0	0	1	0	13:00:00	0	0	1	0	1	1	0
4:00:00	0	0	0	0	0	0	0	13:30:00	1	0	1	0	1	1	2
4:30:00	0	0	0	0	0	0	0	14:00:00	0	0	0	1	0	0	1
5:00:00	5	5	2	5	5	4	4	14:30:00	0	0	2	0	2	0	0
5:30:00	4	5	4	5	4	3	2	15:00:00	0	1	0	0	1	1	0
6:00:00	2	3	2	1	2	4	4	15:30:00	0	1	2	1	0	3	0
6:30:00	2	5	6	2	3	6	5	16:00:00	0	1	2	0	1	1	1
7:00:00	4	2	2	0	3	2	5	16:30:00	2	0	3	1	1	2	1
7:30:00	1	0	0	3	1	2	2	17:00:00	3	0	3	0	0	1	1
8:00:00	1	3	1	1	0	1	1	17:30:00	1	3	0	3	4	1	5
8:30:00	0	1	1	2	2	1	1	18:00:00	4	0	2	0	0	4	1
								18:30:00	5	3	2	2	1	4	2
								19:00:00	1	4	1	1	2	2	1
								19:30:00	3	2	5	4	1	1	1
								20:00:00	6	2	4	3	2	4	3
								20:30:00	0	1	2	2	2	3	2
								21:00:00	0	1	2	0	0	0	2
								21:30:00	5	2	1	0	1	1	3
								22:00:00	2	2	1	3	2	0	1
								22:30:00	2	2	2	0	3	0	0
								23:00:00	0	1	2	1	1	0	2
								23:30:00	1	0	2	0	1	3	1

Tableau 15 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de plage

## Résultats obtenus pour l'analyse de l'activité de la salle de bain

Ceci fait référence à la partie suivante du rapport : 2.3.4.5 *Activité de la salle de bain*

*Machine à laver :*

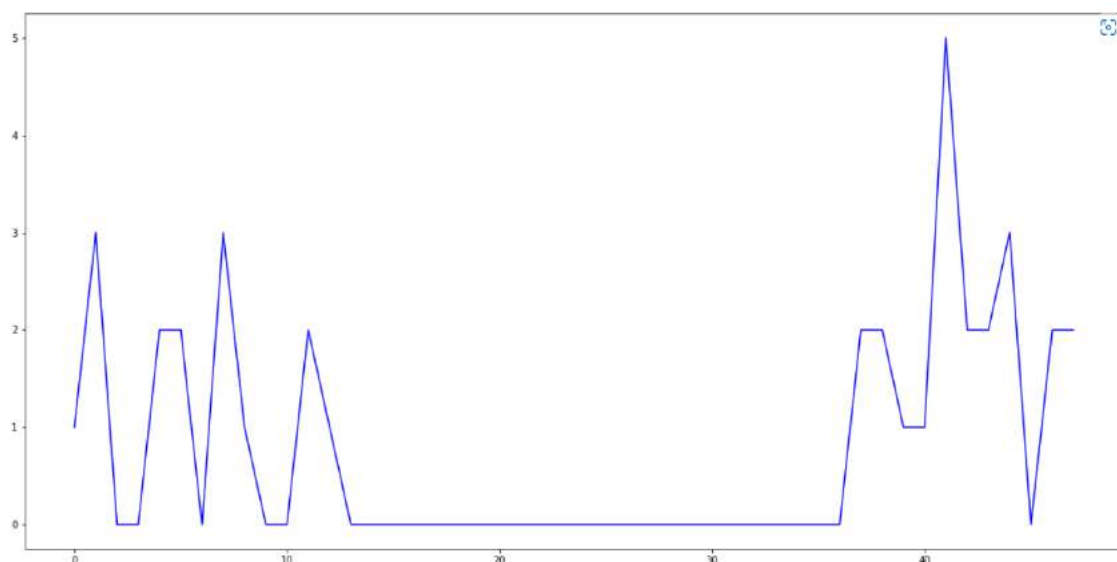


Figure 25 : Visualisation de la fréquence de l'activité correspondant au début de l'activité

	créneau	durée_moy	fréquence								
				16	8:00:00	NaN	0				
0	0:00:00	NaN	0	17	8:30:00	NaN	0	33	16:30:00	NaN	0
1	0:30:00	NaN	0	18	9:00:00	NaN	0	34	17:00:00	0:15:01	2
2	1:00:00	NaN	0	19	9:30:00	0:15:00	2	35	17:30:00	0:27:31	2
3	1:30:00	NaN	0	20	10:00:00	0:25:00	2	36	18:00:00	NaN	0
4	2:00:00	NaN	0	21	10:30:00	0:11:02	1	37	18:30:00	0:24:19	3
5	2:30:00	NaN	0	22	11:00:00	0:15:01	1	38	19:00:00	0:29:01	1
6	3:00:00	NaN	0	23	11:30:00	0:18:37	5	39	19:30:00	NaN	0
7	3:30:00	NaN	0	24	12:00:00	0:23:21	3	40	20:00:00	NaN	0
8	4:00:00	NaN	0	25	12:30:00	0:23:38	2	41	20:30:00	0:15:31	2
9	4:30:00	NaN	0	26	13:00:00	0:29:40	3	42	21:00:00	0:24:01	1
10	5:00:00	NaN	0	27	13:30:00	NaN	0	43	21:30:00	NaN	0
11	5:30:00	NaN	0	28	14:00:00	0:20:30	2	44	22:00:00	NaN	0
12	6:00:00	NaN	0	29	14:30:00	0:17:01	1	45	22:30:00	NaN	0
13	6:30:00	NaN	0	30	15:00:00	0:23:40	3	46	23:00:00	NaN	0
14	7:00:00	NaN	0	31	15:30:00	0:28:41	3	47	23:30:00	NaN	0
15	7:30:00	NaN	0	32	16:00:00	NaN	0				

Tableau 16 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim										
								<b>9:00:00</b>	0	0	0	0	0	0	0	0	0
<b>0:00:00</b>	0	0	0	0	0	0	0	<b>9:30:00</b>	1	0	1	0	0	0	0	0	0
<b>0:30:00</b>	0	0	0	0	0	0	0	<b>10:00:00</b>	0	0	0	2	0	0	0	0	0
<b>1:00:00</b>	0	0	0	0	0	0	0	<b>10:30:00</b>	0	1	0	0	0	0	0	0	0
<b>1:30:00</b>	0	0	0	0	0	0	0	<b>11:00:00</b>	0	0	0	1	0	0	0	0	0
<b>2:00:00</b>	0	0	0	0	0	0	0	<b>11:30:00</b>	4	0	1	0	0	0	0	0	0
<b>2:30:00</b>	0	0	0	0	0	0	0	<b>12:00:00</b>	2	0	0	0	0	0	0	1	0
<b>3:00:00</b>	0	0	0	0	0	0	0	<b>12:30:00</b>	0	0	0	1	1	0	0	0	0
<b>3:30:00</b>	0	0	0	0	0	0	0	<b>13:00:00</b>	1	1	0	1	0	0	0	0	0
<b>4:00:00</b>	0	0	0	0	0	0	0	<b>13:30:00</b>	0	0	0	0	0	0	0	0	0
<b>4:30:00</b>	0	0	0	0	0	0	0	<b>14:00:00</b>	0	0	0	0	2	0	0	0	0
<b>5:00:00</b>	0	0	0	0	0	0	0	<b>14:30:00</b>	0	1	0	0	0	0	0	0	0
<b>5:30:00</b>	0	0	0	0	0	0	0	<b>15:00:00</b>	1	0	0	1	0	0	0	1	0
<b>6:00:00</b>	0	0	0	0	0	0	0	<b>15:30:00</b>	1	0	0	1	1	0	0	0	0
<b>6:30:00</b>	0	0	0	0	0	0	0	<b>16:00:00</b>	0	0	0	0	0	0	0	0	0
<b>7:00:00</b>	0	0	0	0	0	0	0	<b>16:30:00</b>	0	0	0	0	0	0	0	0	0
<b>7:30:00</b>	0	0	0	0	0	0	0	<b>17:00:00</b>	1	0	1	0	0	0	0	0	0
<b>8:00:00</b>	0	0	0	0	0	0	0	<b>17:30:00</b>	0	0	0	0	0	0	0	2	0
<b>8:30:00</b>	0	0	0	0	0	0	0	<b>18:00:00</b>	0	0	0	0	0	0	0	0	0
<i>Tableau 17 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de plage</i>								<b>19:00:00</b>	0	0	0	0	0	1	0	0	0
								<b>19:30:00</b>	0	0	0	0	0	0	0	0	0
								<b>20:00:00</b>	0	0	0	0	0	0	0	0	0
								<b>20:30:00</b>	0	0	0	0	0	0	0	2	0
								<b>21:00:00</b>	1	0	0	0	0	0	0	0	0
								<b>21:30:00</b>	0	0	0	0	0	0	0	0	0
								<b>22:00:00</b>	0	0	0	0	0	0	0	0	0
								<b>22:30:00</b>	0	0	0	0	0	0	0	0	0
								<b>23:00:00</b>	0	0	0	0	0	0	0	0	0
								<b>23:30:00</b>	0	0	0	0	0	0	0	0	0

*Lumière de la salle de bain :*

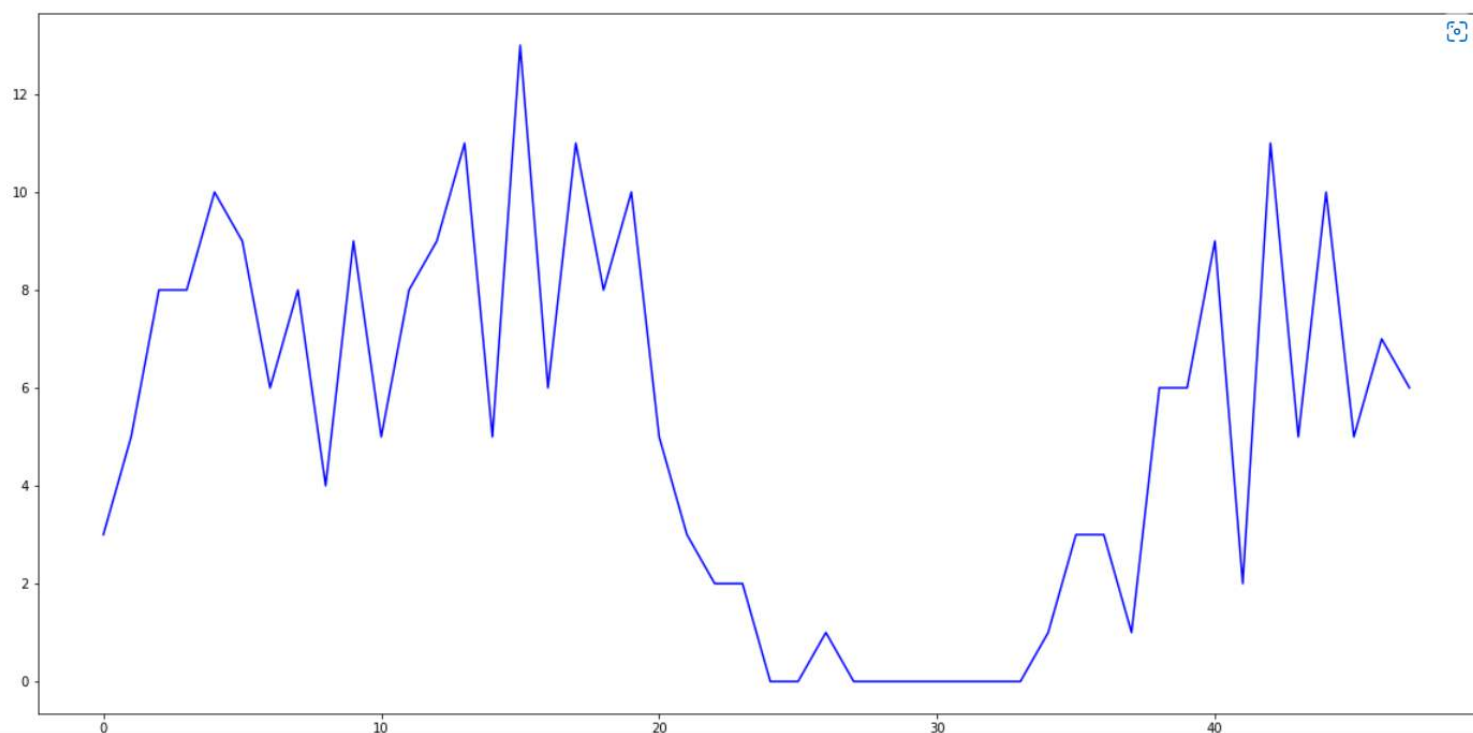


Figure 26 : Visualisation de la fréquence de l'activité correspondant au début de l'activité

	créneau	durée_moy	fréquence								
				15	7:30:00	NaN	0	32	16:00:00	0:17:25	9
0	0:00:00	NaN	0	16	8:00:00	0:03:00	1	33	16:30:00	0:22:19	8
1	0:30:00	0:04:46	10	17	8:30:00	0:07:44	3	34	17:00:00	0:07:14	12
2	1:00:00	0:07:49	7	18	9:00:00	0:07:57	3	35	17:30:00	0:17:14	9
3	1:30:00	0:08:43	3	19	9:30:00	0:08:00	1	36	18:00:00	0:14:32	6
4	2:00:00	0:10:00	2	20	10:00:00	0:15:39	7	37	18:30:00	0:24:52	8
5	2:30:00	0:03:47	2	21	10:30:00	0:12:07	6	38	19:00:00	0:16:27	4
6	3:00:00	NaN	0	22	11:00:00	0:16:46	9	39	19:30:00	0:29:43	9
7	3:30:00	NaN	0	23	11:30:00	0:05:00	2	40	20:00:00	1:18:40	6
8	4:00:00	0:03:00	1	24	12:00:00	0:15:36	11	41	20:30:00	0:22:14	8
9	4:30:00	NaN	0	25	12:30:00	0:21:36	5	42	21:00:00	0:35:09	9
10	5:00:00	NaN	0	26	13:00:00	0:07:18	10	43	21:30:00	0:35:37	11
11	5:30:00	NaN	0	27	13:30:00	0:08:07	5	44	22:00:00	0:27:54	7
12	6:00:00	NaN	0	28	14:00:00	0:06:34	9	45	22:30:00	0:11:14	13
13	6:30:00	NaN	0	29	14:30:00	0:10:29	6	46	23:00:00	0:15:55	7
14	7:00:00	NaN	0	30	15:00:00	0:03:48	5	47	23:30:00	0:29:29	11
				31	15:30:00	0:11:59	5				

Tableau 18 : Table des moyennes de durée et fréquence de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de la plage

	Lun	Mar	Merc	Jeu	Ven	Sam	Dim
0:00:00	0	0	0	0	0	0	0
0:30:00	1	2	1	1	1	3	1
1:00:00	0	1	0	1	1	2	2
1:30:00	0	1	0	0	0	2	0
2:00:00	0	0	0	0	0	1	1
2:30:00	0	0	0	0	0	1	1
3:00:00	0	0	0	0	0	0	0
3:30:00	0	0	0	0	0	0	0
4:00:00	0	0	0	1	0	0	0
4:30:00	0	0	0	0	0	0	0
5:00:00	0	0	0	0	0	0	0
5:30:00	0	0	0	0	0	0	0
6:00:00	0	0	0	0	0	0	0
6:30:00	0	0	0	0	0	0	0
7:00:00	0	0	0	0	0	0	0
7:30:00	0	0	0	0	0	0	0
8:00:00	0	0	0	0	0	1	0
8:30:00	1	0	2	0	0	0	0

Tableau 19 : Calendrier des fréquences de participation à l'activité selon l'appartenance de l'activité à un créneau pour l'horaire de début de plage

9:00:00	1	0	0	1	0	0	1
9:30:00	1	0	0	0	0	0	0
10:00:00	1	3	2	0	0	0	1
10:30:00	2	2	1	0	0	0	1
11:00:00	3	0	1	1	3	0	1
11:30:00	0	0	1	0	0	1	0
12:00:00	1	1	0	1	0	4	4
12:30:00	0	1	1	3	0	0	0
13:00:00	2	1	1	2	3	0	1
13:30:00	2	1	0	0	1	1	0
14:00:00	3	0	0	0	3	1	2
14:30:00	2	1	1	1	0	0	1
15:00:00	1	1	0	0	2	0	1
15:30:00	1	0	1	2	0	1	0
16:00:00	4	0	3	0	0	1	1
16:30:00	1	0	0	2	1	2	2
17:00:00	4	1	2	1	1	2	1
17:30:00	1	2	4	0	0	0	2
18:00:00	1	0	1	0	3	1	0
18:30:00	1	2	0	0	2	1	2
19:00:00	0	0	2	0	1	1	0
19:30:00	4	1	0	0	2	1	1
20:00:00	1	1	1	1	0	1	1
20:30:00	3	2	0	0	1	0	2
21:00:00	1	1	1	1	0	1	4
21:30:00	1	3	2	1	0	1	3
22:00:00	1	2	1	2	0	0	1
22:30:00	1	3	2	0	1	3	3
23:00:00	0	1	2	1	0	0	3
23:30:00	1	2	5	0	1	0	2



## Résultats obtenus pour l'analyse de l'apparition de la lettre suivante

Ceci fait référence à la partie suivante du rapport : 2.4.1.1 Analyse sur l'apparition de la lettre suivante

Lettre	Occurrence	%	Lettre	Occurrence	%
a	5199	27,60%	i	233	1,26%
b	1197	6,45%	j	37	0,20%
c	2129	11,48%	k	191	1,03%
d	2489	13,42%	l	74	0,40%
e	6067	33,13%	m	14	0,08%
f	368	1,98%	n	38	0,20%
g	309	1,67%	o	29	0,16%
h	156	0,84%	p	19	0,10%

Tableau 20 : Table des occurrences d'apparition des termes dans la chaîne

Lettre	Occurrence	%	Lettre	Occurrence	%
a	4210	31,13 %	i	233	1,72%
b	1144	8,46%	j	37	0,27%
c	1982	14,65%	k	191	1,41%
d	2271	16,79%	l	74	0,55%
e	4988	37,52%	m	14	0,10%
f	361	2,67%	n	38	0,28%
g	308	2,28%	o	29	0,21%
h	156	1,15%	p	19	0,14%

Tableau 21 : Table des occurrences d'apparition des termes dans la chaîne (doublons de blocs traités)

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
a		0,05	0,11	0,06	0,73	0,007	0,004	0,006	0,02	0,003	0,009	0,003	0,0004	0,001	0,001	
b	0,18		0,23	0,24	0,21	0,03	0,06	0,02	0,007	0,001	0,002	0,001	0,001	0,002	0,001	0,001
c	0,26	0,15		0,23	0,20	0,08	0,03	0,008	0,01	0,005	0,006	0,005	0,001	0,006	0,003	
d	0,12	0,13	0,23		0,35	0,04	0,04	0,03	0,03	0,001	0,03	0,006	0,001	0,003	0,001	0,006
e	0,62	0,05	0,09	0,18		0,01	0,01	0,004	0,01	0,001	0,005	0,004	0,0003	0,001	0,001	0,001
f	0,1	0,08	0,5	0,16	0,1		0,025	0,01		0,003	0,003	0,01	0,006	0,003	0,006	
g	0,05	0,27	0,21	0,22	0,22	0,02		0,02	0,003							
h	0,12	0,14	0,27	0,3	0,13	0,02	0,02		0,01							
i	0,26	0,02	0,06	0,22	0,39		0,005	0,005			0,03	0,005				0,005
j	0,54	0,14	0,18		0,11							0,04				
k	0,27	0,03	0,1	0,38	0,21	0,14		0,007					0,007			
l	0,25		0,16	0,35	0,15	0,02	0,02		0,04		0,02					
m	0,18	0,18	0,45	0,09		0,09										
n	0,19	0,1	0,32	0,13	0,06	0,16									0,03	
o	0,13		0,56	0,04	0,09	0,13					0,04					
p	0,07	0,07	0,1	0,27	0,2	0,2		0,07								

Tableau 22 : Table des probabilités d'apparition de la lettre suivante (colonne) pour une lettre choisie (ligne)

N.B. : Tout types de doublons ont été supprimés. Les valeurs surlignées correspondant aux plus élevées

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
a		0,03	0,07	0,04	<b>0,48</b>	0,005	0,002	0,005	0,01	0,002	0,009	0,002	0,0002	0,001	0,001	
b	0,14	0,008	0,19	0,22	0,23	0,02	0,06	0,02	0,008	0,003	0,003	0,003	0,001	0,004	0,001	0,001
c	0,22	0,13	0,03	0,23	0,26	0,07	0,03	0,008	0,02	0,006	0,007	0,008	0,0005	0,005	0,002	
d	0,10	0,11	0,20	0,05	<b>0,33</b>	0,03	0,03	0,03	0,24	0,0005	0,03	0,005	0,001	0,003	0,002	0,006
e	<b>0,61</b>	0,06	0,1	0,2	0,3	0,01	0,016	0,006	0,014	0,0015	0,008	0,004	0,0006	0,001	0,001	0,001
f	0,11	0,08	<b>0,49</b>	0,23	0,16		0,03	0,01	0,003	0,003	0,003	0,006	0,006	0,003	0,01	
g	0,05	<b>0,27</b>	0,21	0,26	0,26	0,03	0,04	0,02	0,004		0,004					
h	0,12	0,14	0,27	<b>0,34</b>	0,23	0,02	0,04		0,008							
i	0,26	0,02	0,06	0,22	<b>0,51</b>		0,005	0,005			0,03	0,01				0,005
j	<b>0,54</b>	0,14	0,18		0,32		0,04					0,04				
k	0,26	0,03	0,1	<b>0,39</b>	0,38	0,02		0,007			0,02		0,007		0,007	
l	0,25	0,18	0,16	<b>0,38</b>	0,25	0,02	0,02		0,04		0,018	0,018				
m	0,18	0,18	<b>0,45</b>	0,09	0,18	0,09										
n	0,19	0,1	<b>0,35</b>	0,16	0,26	0,19									0,03	
o	0,13		<b>0,61</b>	0,09	0,17	0,13					0,09					
p	0,07	0,07	0,13	0,26	0,33	0,2		0,07								

Tableau 23 : Table des probabilités d'apparition de la lettre suivante (colonne) par rapport à une lettre choisie (ligne), considérant toutes les lettres d'un même bloc

### Résultats obtenus pour l'analyse de l'appartenance d'activités à un bloc

Ceci fait référence à la partie suivante du rapport : 2.4.1.2 Analyse sur l'appartenance d'activités à un bloc

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	/
a		0,02	0,03	0,02	0,31	0,00 3	0,00 1	0,00 2	0,00 4	0,00 1	0,00 3	0,00 2		0,00 1		0,00 02	0,64
b	0,06		0,02	0,16	0,09	0,00 2		0,00 3			0,00 2	0,00 1			0,00 1		0,72
c	0,07	0,01		0,07	0,08		0,01	0,00 1	0,00 05		0,00 9		0,00 2	0,00 1	0,00 1	0,00 1	0,78
d	0,04	0,08	0,06		0,14	0,00 6	0,00 6		0,00 2		0,00 5	0,00 2		0,00 1	0,00 1		0,69
e	0,25	0,02	0,03	0,06		0,00 2	0,00 2	0,00 2	0,00 5	0,00 1	0,00 4	0,00 1		0,00 1		0,00 02	0,65
f	0,04	0,00 6		0,04	0,03		0,01	0,00 3			0,00 8	0,00 3			0,00 3	0,00 3	0,87
g	0,02		0,07	0,04	0,03	0,02					0,00 3	0,00 3					0,83
h	0,06	0,03	0,01		0,05	0,00 6											0,85
i	0,06		0,00 4	0,02	0,1						0,02	0,00 4					0,80
j	0,11				0,16												0,76
k	0,07	0,01	0,09	0,06	0,09	0,01 6	0,00 5		0,02						0,01		0,71
l	0,1	0,01		0,05	0,05	0,01	0,01		0,01								0,74
m			0,21														0,79
n	0,08		0,05	0,05	0,08												0,82
o		0,03	0,07	0,07		0,03					0,07						0,79
p	0,05		0,1		0,05	0,05											0,79

Tableau 24 : Table des probabilités d'appartenance d'une lettre (colonne) au sein d'un même bloc qu'une lettre choisie (ligne)

N.B. : en surligné vert, les valeurs les plus significatives (élevées) et jaune, un peu moins significatives mais démarquées

## Résultats obtenus pour l'analyse de l'apparition de la lettre suivante sans les capteurs de mouvement

Ceci fait référence à la partie suivante du rapport : 2.4.1.3 *Analyse sur l'apparition de la lettre suivante sans les capteurs de mouvement*

Lettre	Occurrence	%	Lettre	Occurrence	%
f	221	23%	k	156	16,23%
g	161	16,75%	l	40	4,16%
h	130	13,53%	m	13	1,35%
i	148	15,4%	n	36	3,75%
j	30	3,12%	o	28	2,91%
			p	19	1,98%

Tableau 25 : Table des occurrences d'apparition des termes dans la chaîne sans capteurs de mouvement

	f	g	h	i	j	k	l	m	n	o	p
f		0,34	0,2	0,11	0,03	0,1	0,07	0,01	0,03	0,03	0,03
g	0,46	0,02	0,29	0,12	0,04	0,06	0,02	0,006	0,006	0,006	0,006
h	0,36	0,29		0,16	0,04	0,1	0,02	0,008	0,02	0,02	0,02
i	0,08	0,1	0,08		0,03	0,55	0,05	0,007	0,03	0,01	0,003
j	0,17	0,1	0,13	0,4		0,13	0,03				0,03
k	0,26	0,09	0,09	0,35	0,01	0,03	0,03	0,03	0,08	0,05	0,02
l	0,22	0,16	0,11	0,22	0,05	0,14			0,08	0,03	
m	0,31	0,08		0,23		0,31			0,08		
n	0,31	0,06	0,08	0,11	0,03	0,08	0,14	0,05		0,17	
o	0,4	0,04		0,04		0,28	0,04		0,16	0,04	0,04
p	0,5	0,17	0,06	0,06		0,11		0,06	0,06		

Tableau 26 : Table des probabilités d'apparition de la lettre suivante (colonne), selon une lettre choisie (ligne) sans les capteurs de mouvement

N.B. : en surligné vert, les valeurs significatives élevées, en jaune les valeurs d'une signification secondaire

## Résultats obtenus pour l'analyse de l'apparition de la lettre suivante avec un changement d'activité

Ceci fait référence à la partie suivante du rapport : 2.4.1.4 *Analyse sur l'apparition de la lettre suivante avec un changement d'activité*

Lettre	Occurrence	%	Lettre	Occurrence	%
q	235	25,19%	k	155	16,61%
g	158	16,93%	l	41	4,39%
h	129	13,83%	n	36	3,86%
i	148	15,86%	p	19	2,04%
j	30	3,22%			

Tableau 27 : Table des occurrences des termes dans la chaîne sans les capteurs de mouvement et pour la réunion de capteurs sous l'activité « faire un repas » (q)

	q	g	h	i	j	k	l	n	p
q	0,009	0,35	0,2	0,13	0,03	0,16	0,07	0,04	0,03
g	0,45		0,28	0,11	0,04	0,06	0,02	0,006	0,006
h	0,4	0,27		0,16	0,04	0,1	0,02	0,02	0,02
i	0,11	0,1	0,08		0,03	0,54	0,05	0,03	0,03
j	0,2	0,1	0,13	0,4		0,13	0,03		0,03
k	0,31	0,08	0,09	0,35	0,02	0,03	0,03	0,08	0,01
l	0,31	0,13	0,1	0,21	0,05	0,13		0,08	
n	0,53	0,06	0,08	0,11	0,03	0,08	0,14		
p	0,53	0,16	0,05	0,05		0,11		0,05	

Tableau 28 : Table des probabilités d'apparition de la lettre suivante (colonne) selon une lettre choisie (ligne) sans les capteurs de mouvement et pour la réunion de capteurs sous l'activité « faire un repas » (q)

N.B. : en surligné vert, les valeurs significatives élevées, en jaune, les valeurs significatives mais moins importantes.

## Résultats obtenus pour l'analyse d'occurrence de séquences de caractères

Ceci fait référence à la partie suivante du rapport : 2.4.1.5 Analyse d'occurrence de séquences de caractères

Nb	Expression	Proba	Nb	Expression	Proba	Nb	Expression	Proba	
2	ae	0,217	3	aea	0,189	4	aeae	0,17	
	ea	0,216		eae	0,186		eaea	0,17	
	de	0,059		ded	0,038		dede	0,031	
	ed	0,054		ede	0,037		eded	0,029	
	cd	0,028		aca	0,011		aeac	0,006	
	ec	0,028		aec	0,011		caea	0,006	
	ce	0,027		cae	0,011		aeca	0,005	
	dc	0,027		dcd	0,011		aead	0,005	
	ca	0,025		cdc	0,01		cdcd	0,005	
	ac	0,024		eac	0,01		ecae	0,005	
	bd	0,021		eca	0,009		eaec	0,005	
	db	0,015		cea	0,009	<b>Commentaire :</b> Mêmes remarques que pour 2 caractères. De plus en plus de termes qui ne paraissent pas porter de sens.			
	ad	0,014		ace	0,008				
	be	0,014		cac	0,008				
	cb	0,014		ead	0,008				
	bc	0,013		ece	0,008				
	cf	0,013		cec	0,008	5	aeaea	0,158	
	da	0,013		dbd	0,007		eaeae	0,155	
	ab	0,011		cfc	0,007		deded	0,026	
	ba	0,009		dae	0,007		edede	0,026	
	bg	0,005		bdc	0,006	<b>Commentaire :</b> Mêmes remarques			
	dh	0,005		cde	0,006				
	ei	0,005		dea	0,006				
	ie	0,005		dec	0,006				
	<b>Commentaire :</b> Ecart flagrant avec deux expressions <i>largement</i> dominantes mais qui ne			eab	0,006				
				aed	0,006	6	aeaeae	0,147	
				abe	0,005		eaeaea	0,145	
				ada	0,005		dedede	0,024	

reflètent pas des séquences <i>sensées</i> .			ade	0,005		ededed	0,023
			<b>aeb</b>	0,005	<b>Commentaire :</b> Mêmes remarques		
			bdb	0,005			
			bde	0,005			
			<b>bea</b>	0,005			
			<b>cbd</b>	0,005			
			<b>cdb</b>	0,005			
			ced	0,005			
			ebe	0,005			
			edc	0,005			
			<b>Commentaire :</b> On remarque qu'on obtient des chaînes uniquement avec les caractères correspondants aux capteurs de mouvement.				

Tableau 29 : Table des probabilités d'apparition de séquences de caractères selon la longueur de la séquence  
N.B. : en surligné, les expressions qui ne paraissent pas sensées, a priori

Résultats obtenus pour l'analyse d'occurrence de séquences de caractères sans les capteurs de mouvement

Ceci fait référence à la partie suivante du rapport : 2.4.1.6 Analyse d'occurrence de séquences de caractères sans les capteurs de mouvement

Nb	Expression	Proba	Nb	Expression	Proba	Nb	Expression	Proba
2	ik	0,087	3	<b>gfg</b>	0,05	4	<b>fgfg</b>	0,033
	fg	0,081		<b>fgf</b>	0,049		<b>gfgf</b>	0,032
	gf	0,077		<b>kik</b>	0,04		ghgf	0,02
	ki	0,054		fgh	0,027		gfgf	0,016
	hf	0,048		iki	0,024		<b>ikik</b>	0,015
	fh	0,047		<b>fhf</b>	0,023		fghg	0,014
	gh	0,045		ikf	0,022		<b>fhfh</b>	0,013
	kf	0,04		<b>hfh</b>	0,019		hgfg	0,013
	hg	0,036		hgf	0,019		<b>hfhf</b>	0,012
	fk	0,025		ghf	0,017		fghf	0,011
	fi	0,024		<b>ghg</b>	0,017		<b>kiki</b>	0,011
	hi	0,021		gfh	0,016		<b>fgfh</b>	0,01
	gi	0,018		hfg	0,015		<b>fhgf</b>	0,009
	fl	0,016		fhg	0,014		hfgf	0,009
	ig	0,015		ikn	0,013		kikf	0,009
	kg	0,014		fik	0,011		gfhf	0,008
	kh	0,014		fki	0,011		gfhg	0,008
	ih	0,013		hik	0,01		fkik	0,006
	kn	0,013		<b>kfi</b>	0,01		hfgh	0,006
	if	0,012		<b>hgh</b>	0,009		ikfg	0,006
	ji	0,012		ikh	0,009		ikfh	0,006
	hk	0,011		lg	0,007		kfik	0,006
	nf	0,011		fkf	0,006		fhgh	0,005
	<b>of</b>	0,01		hfk	0,006		Fiki	0,005

	gk	0,009		fn	0,006		ghfh	0,005
	pf	0,009		no	0,006		hfhg	0,005
	il	0,008		fif	0,005		hghf	0,005
	ko	0,008		gfk	0,005		hkik	0,005
	lf	0,008		ghi	0,005		ikhi	0,005
	li	0,008		hih	0,005	iknf	0,005	
	ok	0,008		hki	0,005	kikh	0,005	
	fj	0,007		igi	0,005	kikn	0,005	
	fo	0,007		ikg	0,005	okik	0,005	
	fp	0,007		kfo	0,005	Commentaire :		
	gj	0,007		kgf	0,005			
	lg	0,007	Commentaire : Chaînes avec caractères répétitifs					
	fn	0,006						
	no	0,006						
	hj	0,005						
	jf	0,005						
	kl	0,005						
	lh	0,005						
	lk	0,005						

Commentaire : Pas même un pourcentage au-dessus de 10%.			
--	--	--	--

6	fgfgfg	0,015	5	gfgfg	0,024
	gfgfgf	0,013		fgfgf	0,02
	gfgfgh	0,011		fgfgh	0,013
	hgfgfg	0,009		hgfgf	0,011
	fgfghf	0,008		fhfhf	0,009
	fghgfg	0,007		gfgfh	0,009
	gfghfg	0,007		hfgfg	0,008
	hfgfgf	0,007		hfhfh	0,008
	fhfhfh	0,006		fghfg	0,007
	ghgfgf	0,006		gfgfh	0,007
	fgfgfh	0,005		gfghg	0,007
	fghghg	0,005		ghgfg	0,007
	gfgfhg	0,005		fgfhg	0,006
	gfghgf	0,005		gfhgf	0,006
	gfhfhf	0,005		ghfgh	0,006
	ghfgfg	0,005		ikikf	0,006
	ghfghg	0,005		kikik	0,006
				gfhfh	0,005
				ghfgf	0,005
		hfgfhg	0,005		

Commentaire :			
---------------	--	--	--


Commentaire :		
---------------	--	--

Tableau 30 : Table des probabilités d'apparition de séquences de caractères selon la longueur de la séquence sans les capteurs de mouvement

## Bibliographie et sources

[1] CHIMAMIWA Gibson, ALIREZAIE Marjan, PECORA Federico, LOUFTI Amy, "Multi-sensor dataset of human activities in a smart home environment", Mendeley Data, November 2020.

Commentaire : jeu de données exploitable sur lequel nous avons travaillé pour ce stage

Source : communication de l'équipe

Mots-clé : enregistrement de capteurs, données dans un appartement intelligent, reconnaissance d'activité.

[2] DEROCHÉ Léna, OLISLAGERS Antoine, SABOURDY Anna, « Mission R&D N°8 : Construction automatique de scénarios utilisateurs dans les environnements intelligents », Avril 2022

Commentaire : rapport de mission d'un groupe d'élèves de *IMT Mines Alès* sur le projet HUT. Nous avons essentiellement poursuivi leur travail en faisant des analyses plus précises.

Source : communication de l'équipe

Mots-clés : scénario utilisateur, prédiction de scénario.

[3] DESSE Olivier, « Stage M1 : Projet HUT », Juillet 2021

Commentaire : rapport de stage d'un élève de *IMT Mines Alès* sur le projet HUT

Source : communication de l'équipe

Mots-clés : algorithme prédictif, LZ78, LeZi, MLP, ARIMA

[4] OUKRICH Nadia, « Daily Human Activity Recognition in Smart Home based on Feature Selection, Neural Network and Load Signature of Appliances », Technology for Human Learning, Université Mohamed V, 2019.

Commentaire : travail similaire basé sur la reconnaissance d'activité dans une maison intelligente.

Source : HAL Archives Ouvertes

Mots-clés : reconnaissance d'activité, maison connectée.

## Glossaire

« smart home » (maison intelligente) : désigne une habitation équipée d'appareils connectés, voire capable d'effectuer des opérations automatiquement

Reconnaissance d'activité : reconnaissance des actions et des objectifs d'un ou de plusieurs agents à partir d'une série d'observations sur les actions des agents et des conditions environnementales

Prédiction d'activités (ou de scénario) : Action d'annoncer à l'avance un événement par calcul, par raisonnement, par induction. Ici, la prédiction de scénario consiste à prédire les actions des utilisateurs dans l'environnement connecté à partir des actions précédentes.

Algorithme de prédiction : Modèle mathématique conçu pour anticiper les résultats futurs les plus certains en tenant de résultats survenus précédemment.

# Synthèse

Mots clés : maison connectée ; reconnaissance d'activité ; prédiction d'activité ; recherche ; capteur ; motif ; appareil connecté

J'ai passé 13 semaines au sein du laboratoire de recherche CERIS de mon école *IMT Mines Alès* à l'occasion de mon stage de deuxième année. J'ai travaillé au sein de l'équipe sur le projet HUT des maisons connectées, encadrée par mon professeur *M. Sylvain Vauttier*.

L'objectif de ce stage était d'étudier un jeu de données pour ainsi faire de la prédiction d'activités. Avec mon collègue, nous avons donc mené différentes analyses pour étudier le comportement des capteurs, ou des combinaisons de plusieurs capteurs, pour déceler des motifs ou des habitudes de comportement dans l'utilisation des différents appareils connectés. Nous avons aussi étudié l'enchaînement des activités en faisant des analyses probabilistes sur des chaînes de caractères.

Un autre objectif de ce stage était de travailler sur l'implémentation de jumeaux numériques pour ce projet, mais malheureusement nous n'avons pas pu approfondir cette idée.

Cette expérience m'a permis de découvrir plus amplement le domaine de la recherche, notamment avec ces difficultés. Cela m'a permis de mieux envisager le monde professionnel dans mon domaine, et je ressors très satisfaite de cette expérience

## Abstract

Key words: smart home ; activity recognition ; activity prediction ; research ; sensor ; pattern ; smart device

I spent 13 weeks in the research lab CERIS at my school *IMT Mines Alès* for my second-year internship. I work within the team on the HUT project on smart homes, supervised by my professor *Mr. Sylvain Vauttier*.

The goal of this internship was to study a dataset to make activity prediction. With my colleague, we led different analysis in order to study the behavior of the sensors, or with a combination of several sensors, to identify patterns or behavioral habits in the use of the different smart devices. We also studied activities sequences by doing probabilistic analysis on character sequences.

Another goal of this internship was to work on the implementation of digital twins for this project, but unfortunately, we hadn't got the chance to go far with this idea.

This experience allowed me to discover more broadly the research field with its difficulties. It gave me another view of the professional experience in my field, and I feel very satisfied of this experience.