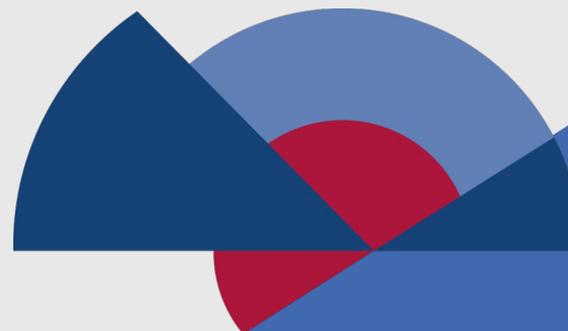


Comment utiliser les para-données pour améliorer son questionnaire en vue d'une collecte auto- administrée ? Le cas de l'enquête Logement



- 1 LES PARADONNÉES DE LA NOUVELLE FILIÈRE**
- 2 EXEMPLES D'UTILISATION DES PARADONNÉES
DU PILOTE DE L'ENL**
- 3 TRAVAUX À VENIR SUR LES PARADONNÉES**

01

LES PARADONNÉES DE LA NOUVELLE FILIÈRE

Les paradonnées sont des enregistrements horodatés de toutes les actions d'une unité répondante par unité répondante et session de réponse.

Les actions enregistrées sont :

- entrée dans le questionnaire,
- coche de chaque réponse avec la valeur ad hoc,
- saisie dans un champ texte ou numérique avec la valeur correspondante,
- champ de réponse laissé à vide,
- clic sur le bouton de retour arrière,
- clic sur le bouton suivant,
- sortie du questionnaire.



Les paradonnées sont disponibles pour tous les questionnaires auto-administrés par internet.



Exemple des parodonnées brutes de réponses courtes au questionnaire pilote de l'ENL.

CADR

Confirmez-vous que vous habitez toujours à l'adresse suivante : \$ADRESSE\$?

- 1 - Oui, j'habite toujours à cette adresse et l'adresse est correcte et complète
- 2 - Oui, j'habite toujours à cette adresse, mais l'adresse est incorrecte ou incomplète
- 3 - Non, j'ai déménagé et habite à une autre adresse
- 4 - Non, je n'ai jamais habité à cette adresse

INDNVOCC

Connaissez-vous le nom de la ou des personnes qui vous ont succédé dans le logement situé à l'adresse suivante : \$ADRESSE\$?

- 1 - Oui
- 2 - Non

idSurveyUnit	typeParadataObject	idOrchestrator	idQuestionnaire	userAgent	page	idSession	type	idParadataObject	timestamp	responseName	value	oldValue	newValue	battery
1	PL1100000101	session	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	NA	593d88e6-3f11-49a2-9727-92a8e13fb39f	session-started	init-session	1645807624027	NA	nan	nan	nan
2	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	NA	593d88e6-3f11-49a2-9727-92a8e13fb39f	orchestrator-create	init-orchestrator-collect	1645807624458	NA	nan	nan	nan
3	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	welcomePage	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	next-button-orchestrator-collect	1645807637176	NA	nan	nan	nan
4	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	1	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	next-button-orchestrator-collect	1645807674779	NA	nan	nan	nan
5	PL1100000101	lunatic-input	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	1	593d88e6-3f11-49a2-9727-92a8e13fb39f	selection	radio-kb9hlpdc-3	1645807679901	CADR	3	nan	nan
6	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	2	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	next-button-orchestrator-collect	1645807684715	NA	nan	nan	nan
7	PL1100000101	lunatic-input	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	2	593d88e6-3f11-49a2-9727-92a8e13fb39f	selection	radio-kqweh61i-2	1645807690025	INDNVOCC	2	nan	nan
8	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	4	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	next-button-orchestrator-collect	1645807691450	NA	nan	nan	nan
9	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	184	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	next-button-orchestrator-collect	1645807730751	NA	nan	nan	nan
10	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	validationPage	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	validate-button-orchestrator-collect	1645807734894	NA	nan	nan	nan
11	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	validationPage	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	agree-sending-modal-button-orchestrator-collect	1645807741929	NA	nan	nan	nan
12	PL1100000101	orchestrator	orchestrator-collect	LOG2022X01	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWe...	endPage	593d88e6-3f11-49a2-9727-92a8e13fb39f	click	logout-close-button-orchestrator-collect	1645807759129	NA	nan	nan	nan

idSurveyUnit	page	responseName	timestamp_converted	numEvent	page_num	NumSession	temps_event_m	temps_event_m_inf15	temps_event_m_sup15
1	PL1100000101	NA	2022-02-25 18:47:04	1	NA	1	0.000000000	0.000000000	0.000000000
2	PL1100000101	NA	2022-02-25 18:47:04	2	NA	1	0.007183333	0.007183333	0.007183333
3	PL1100000101	welcomePage	2022-02-25 18:47:17	3	NA	1	0.211966667	0.211966667	0.211966667
4	PL1100000101	1	2022-02-25 18:47:54	4	1	1	0.626716667	0.626716667	0.626716667
5	PL1100000101	1	2022-02-25 18:47:59	5	1	1	0.085366667	0.085366667	0.085366667
6	PL1100000101	2	2022-02-25 18:48:04	6	2	1	0.080233333	0.080233333	0.080233333
7	PL1100000101	2	2022-02-25 18:48:10	7	2	1	0.088500000	0.088500000	0.088500000
8	PL1100000101	4	2022-02-25 18:48:11	8	4	1	0.023750000	0.023750000	0.023750000
9	PL1100000101	184	2022-02-25 18:48:50	9	184	1	0.655016667	0.655016667	0.655016667
10	PL1100000101	validationPage	2022-02-25 18:48:54	10	NA	1	0.069050000	0.069050000	0.069050000
11	PL1100000101	validationPage	2022-02-25 18:49:01	11	NA	1	0.117250000	0.117250000	0.117250000
12	PL1100000101	endPage	2022-02-25 18:49:19	12	NA	1	0.286666667	0.286666667	0.286666667

02 EXEMPLES D'UTILISATION DES PARADONNÉES DU PILOTE DE L'ENL

Exemple 1 : Utilisation des paradonnées pour valider la conception de chaque question du questionnaire

À partir des paradonnées, des indicateurs ont été constitués pour connaître :

- la durée par question (minimum, maximum, moyenne, médiane, et somme de la durée sur tous les répondants)
- le nombre de corrections par question (minimum, maximum, moyenne, somme sur tous les répondants, le nombre d'UE ayant effectué au moins une correction, et la part d'UE ayant effectué au moins une correction)
- le nombre de retours arrière par question (minimum, maximum, moyenne, somme sur tous les enquêtés, le nombre d'UE ayant effectué au moins un retour arrière, et la part d'UE ayant effectué au moins un retour arrière).
- le nombre d'UE ayant effacé leur réponse par question (la part d'UE ayant répondu puis effacé leur réponse).



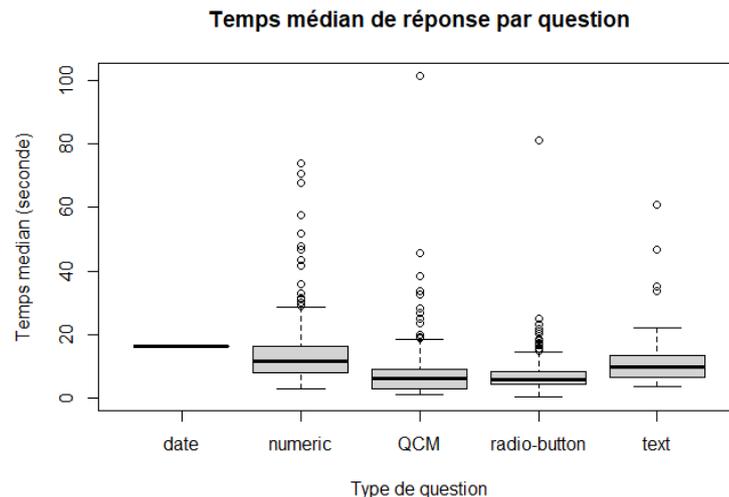
Le fichier DDI des spécifications formelles du questionnaire a été utilisé pour récupérer ou calculer des métadonnées :

- Le nombre de mots par question,
- Le type de question (cases à cocher unique ou multiple, liste déroulante, format numérique ou texte).

Ces métadonnées permettent d'expliquer une partie de la variabilité des indicateurs.



DATA DOCUMENTATION INITIATIVE



Les indicateurs calculés permettent de connaître les questions qui sont les plus compliquées pour les enquêtés.

Exemple : Les variables sur le nombre de salles de bain et de WC

→ 135. Combien avez-vous de salle d'eau ou salle de bain ?

Avec baignoire :	Please, do something...
Avec douche mais sans baignoire :	0
	1
	2
	3
	4
	5
	6
	7
	8
	9

Plusieurs indicateurs sont élevés sur ces questions :

Part d'UE ayant fait un retour arrière :

- 100 % pour la question sur le nombre de salles de bain avec baignoire,
- 78 % pour la question sur le nombre de salles de bain avec douche mais sans baignoire,
- 100 % pour la question sur le nombre de W-C avec baignoire situés dans une pièce indépendante,
- 90 % pour la question sur le nombre de W-C avec baignoire situés dans une autre pièce.

Les enquêtés ont eu des problèmes d'ergonomie sur ces quatre questions qui sont des listes déroulantes au sein d'un tableau.



Les questions sur les salles de bains et WC ont été reformulées sur le modèle suivant:

→ 6. Possédez-vous une salle d'eau ou une salle de bain (pièce contenant une douche ou une baignoire) ?

- Oui
- Non

→ 7. Combien avez-vous de salles d'eau ou de salles de bain ?

→ 8. Votre salle d'eau ou de bain comporte-t-elle une baignoire ?

- Oui
- Non

→ 9. Parmi vos 2 salles d'eau ou de bain, combien comportent une baignoire ?

Exemple 2 : dernière réponse atteinte par session d'un même enquêté



Les parodonnées permettent de savoir quelle est la dernière question atteinte par enquêté et par session de réponse.

Exemple : Sur les questions sur le nombre de salles de bain, 139 UE se sont déconnectées du questionnaire, dont 60 UE qui ne se sont pas reconnectées au questionnaire.

Exemple 3 : recensement des navigateurs et matériels utilisés pour répondre sur internet (exploitation de la variable UserAgent)

1 861 connexions sur 8091 ont été réalisées depuis un smartphone ou une tablette en S1 internet du pilote Logement 2022, dont a priori 350 tablettes et 1 511 smartphones soit respectivement 4,3% et 18,7% des connexions à la S1 internet.



Exemple 4 : effet du support de navigation non pondéré ?



Comment estimez-vous vos conditions de logement actuelles ? Résultats non pondérés	Ensemble	Mobile/ tablette	Ordinateur
Très satisfaisantes	30,3 %	24,8 %	31,9 %
Satisfaisantes	44,5 %	42,9 %	45,0 %
Acceptables	18,3 %	23,0 %	17,0 %
Insuffisantes	4,8 %	6,6 %	4,3 %
Très insuffisantes	2,0 %	2,7 %	1,8 %

Exemple 4 : effet du support de navigation après repondération ?



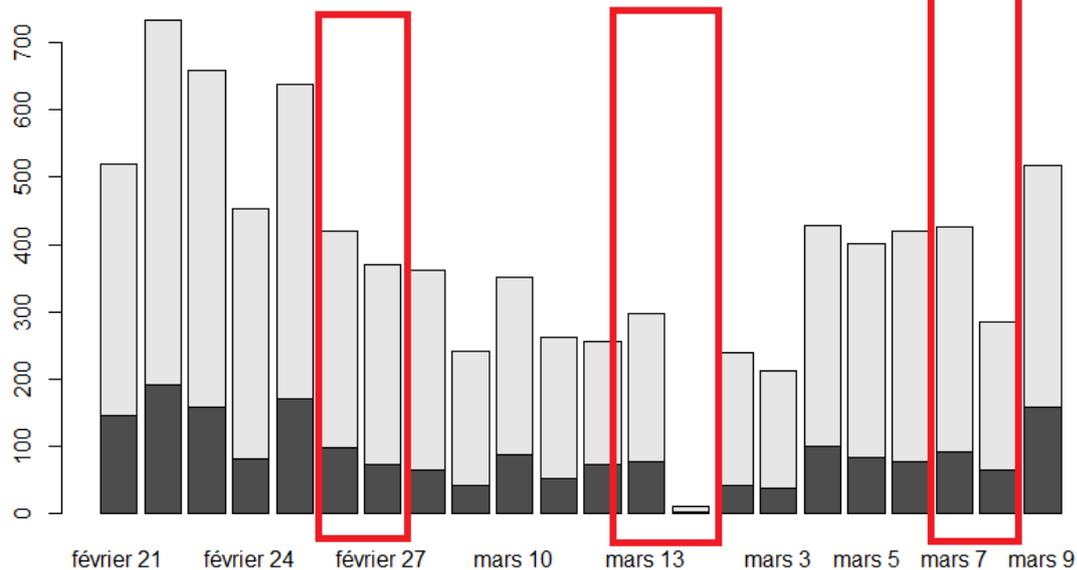
Comment estimez-vous vos conditions de logement actuelles ? Pondérations provisoires (additif)	Ensemble	Mobile/ tablette	Ordinateur
Très satisfaisantes	28,6 %	23,9 %	29,8 %
Satisfaisantes	41,4 %	38,8 %	42,2 %
Acceptables	21,8 %	25,5 %	20,8 %
Insuffisantes	6,2 %	9,5 %	5,3 %
Très insuffisantes	2,0 %	2,3 %	1,9 %

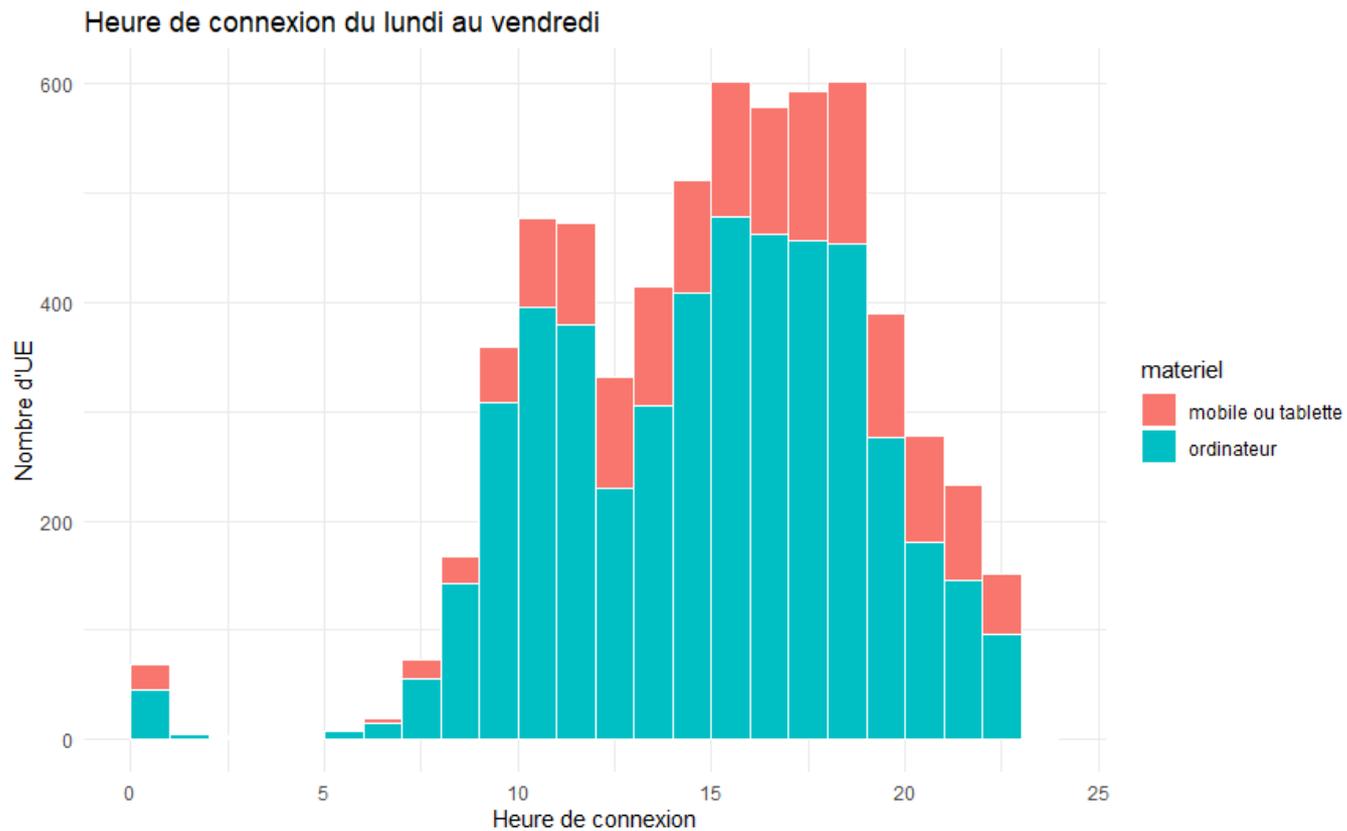
Exemple 5 : utilisation de différents type de support pour un même répondant

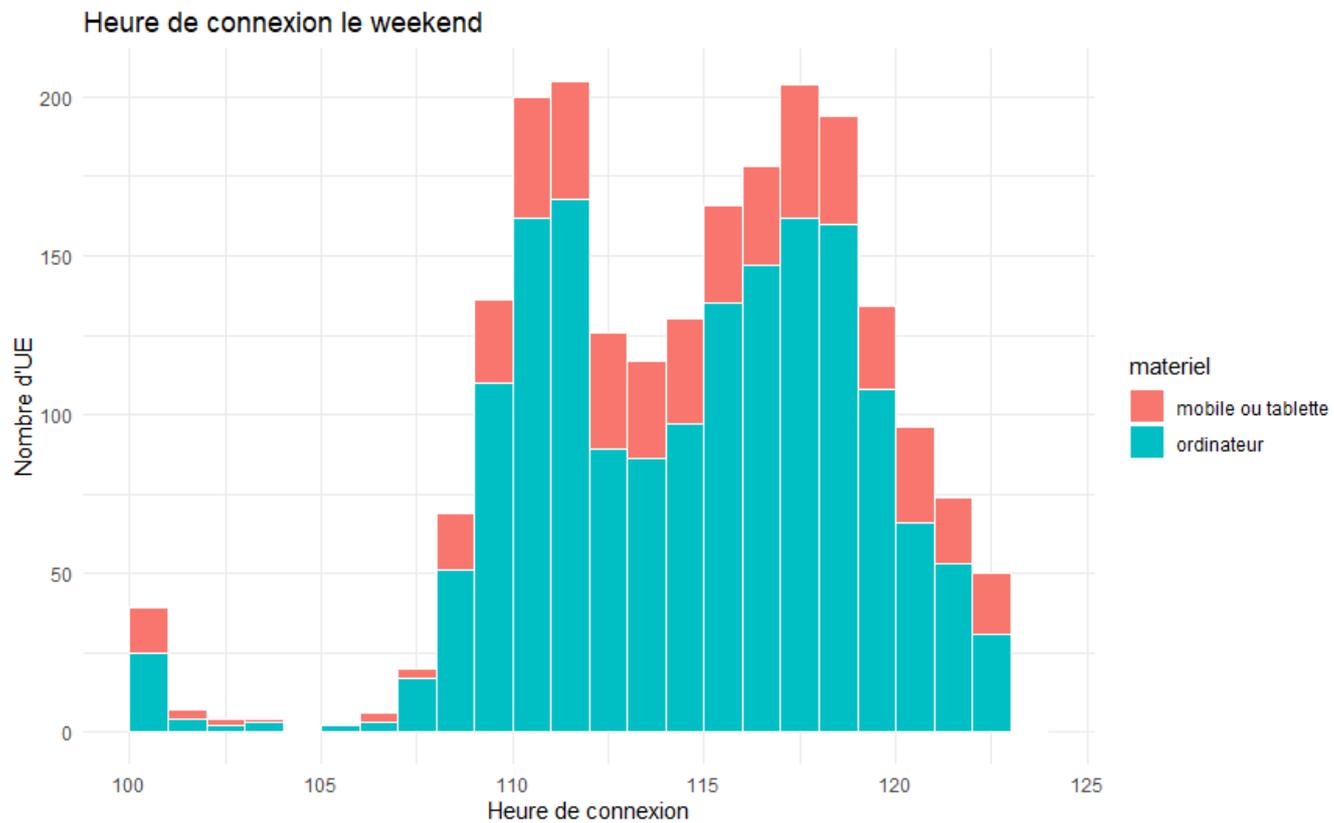


Sur 7 841 UE répondantes en S1 internet au pilote Logement 2022,
6 023 UE ont utilisé seulement un PC,
1 783 ont utilisé seulement un téléphone ou une tablette,
35 ont utilisé les deux types de support.

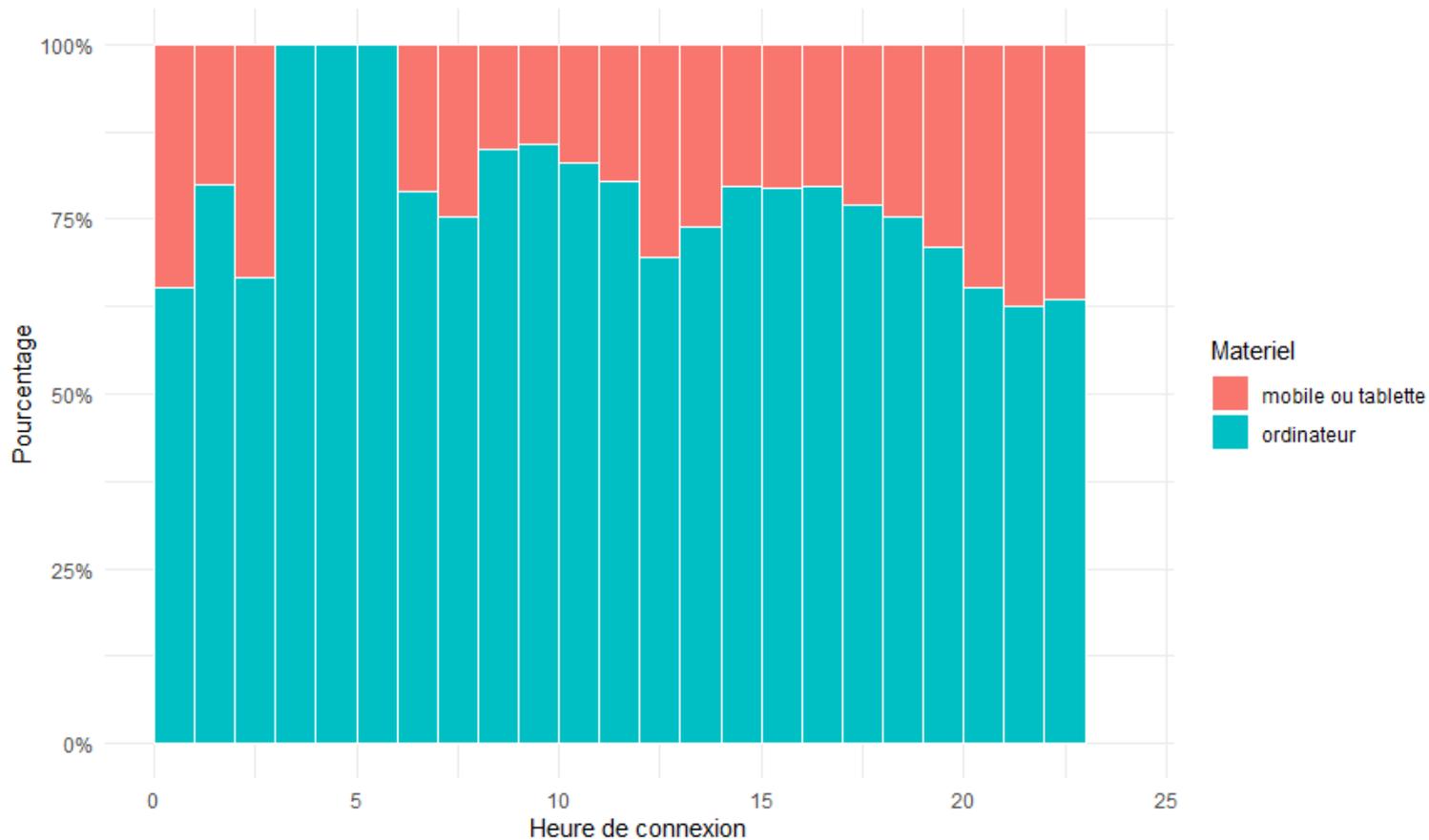
Exemple 6 : Heure et jour de connexion selon le type de support mobile







Heure de connexion du lundi au vendredi



03 TRAVAUX À VENIR SUR LES PARADONNÉES

Les paradonnées peuvent être utilisées pour classier les répondants :

- La durée moyenne par question ;
- Le nombre de retours arrière ;
- Le nombre de changements de réponse ;
- Le nombre de réponses effacées ;
- D'autres indicateurs qui permettraient de savoir si l'enquêté accélère ses réponses au fur et à mesure du questionnaire.



Ces études peuvent servir à :

- détecter des questionnaires mal répondus (réponse trop rapide) ;
- déterminer des parcours type au sein d'un questionnaire.



A vous de jouer :



Programmes python pour constituer une base exploitable de paradonnées (conversion des json en une table parquet ou csv) :

https://gitlab.insee.fr/enquete-nationale-logement/sauvegarde_codes/-/tree/master/pilote2022/paradonnees

Programme R d'exploitation des paradonnées et d'import des métadonnées :

https://gitlab.insee.fr/enquete-nationale-logement/sauvegarde_codes/-/tree/master/paradonnees

https://gitlab.insee.fr/enquete-nationale-logement/sauvegarde_codes/-/tree/master/metadonnees

Paradonnées du Pilote Logement :

X:\HAB-ENL2023-PARADONNEES\Pilote_2022\data

Retrouvez-nous sur

[insee.fr](https://www.insee.fr)



Fabienne Perray-Gibert
Responsable de section
Section Logement

01 87 69 63 64

fabienne.perray-gibert@insee.fr

Lisa Kerboul
Conceptrice d'enquête
Section Logement

01 87 69 63 61

lisa.kerboul@insee.fr