

Le rapport des individus aux questionnaires

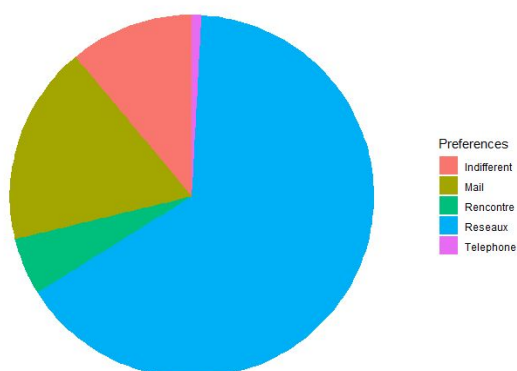
Nous vivons dans une société où chaque activité cherche de plus en plus à prédire le comportement de ses agents (consommateurs, clients, adhérents etc.). Grâce aux TIC et aux outils statistiques et probabilistes, chaque activité cherche à récolter un maximum d'informations sur ses agents pour les utiliser par la suite dans la continuité et le renouvellement de son activité. Ce qui a permis le développement d'une méthode de récolte de données : le questionnaire d'informations. Il peut s'agir de nous, étudiants, dans le cadre de nos études qui devons réaliser une enquête sur un thème, une entreprise qui cherche à développer son activité, ou à avoir des retours sur les consommations de sa clientèle. Les outils de diffusion de ces questionnaires vont ainsi pouvoir varier : un post sur les réseaux sociaux, un mail, un appel téléphonique ou encore un démarchage physique.

100 % de nos 118 enquêtés ont déjà été au moins une fois confronté à un quelconque questionnaire avant d'avoir répondu à notre enquête. L'objectif de notre enquête a pour but de donner un retour sur l'opinion des agents face à ces questionnaires : répondent-ils rigoureusement ? Sont-ils méfiant face à ces questionnaires, face à l'enquêteur ? Quels sont les moyens de diffusion les plus efficaces ? Quels thèmes abordés peuvent poser problème ? Comment inciter les gens à répondre ?

Parmi nos 118 enquêtés, 81% ont entre 18 et 25 ans, 66% sont des femmes et 80% sont étudiants. Mais seulement 48% apprécient répondre à un questionnaire. Nous allons donc essayer de répondre à la problématique suivante en fonction de notre échantillon : **Quel est le rapport des individus face aux questionnaires auxquels ils peuvent être sujets ?**

Les canaux de distribution des questionnaires

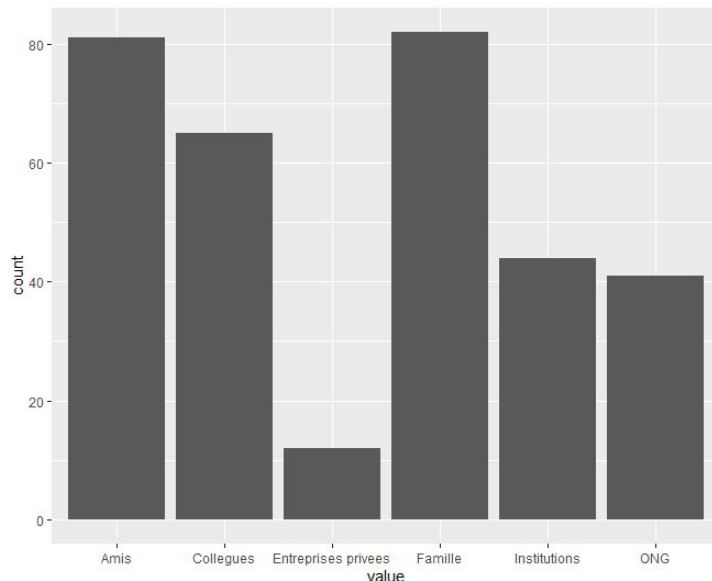
Nous nous sommes intéressés à quatre moyens de distribution dans notre étude, à savoir les réseaux sociaux, les mails, les appels téléphoniques et les rencontres physiques. Parmi nos enquêtés 100% consultent souvent leur boîte mail, 94% utilisent les réseaux sociaux et 60% ont accès à un téléphone fixe. Les individus de notre échantillon ont déjà été soumis à un questionnaire sur les réseaux sociaux à 85%, par mail et par téléphone à 60% et à 40% par un enquêteur physique. Les individus préfèrent répondre aux questionnaires à 65% sur les réseaux sociaux, à 18% par mail, à 5% en rencontrant physiquement la personne et à seulement 1% par téléphone fixe (sachant que 11% sont indifférents). On remarque ainsi que les réseaux sociaux et les mails sont à la fois plus utilisés par les enquêtés et les enquêteurs et qu'ils sont les moyens de soumissions à un questionnaire préférés par la population.



On peut expliquer ces préférences par différents facteurs. Le premier étant le temps maximum qu'un individu est prêt à consacrer à un questionnaire : 46% sont prêts à consacrer 5 minutes maximum et 47% entre 5 et 10 minutes. Ensuite on peut remarquer que parmi les questions préférées des individus on retrouve les QCM à 80%, les questions fermées à 41%, les échelles d'appréciation à 35% et les questions ouvertes à moins de 8%. Ces deux facteurs favorisent les plateformes qui permettent des réponses rapides et simples telles que les réseaux sociaux et les mails.

Les réponses aux questionnaires biaisées par des considérations sociales

Dans notre échantillon, 3 personnes sur 4 répondent en fonction de l'interrogateur tandis que 9 personnes sur 10 choisissent de répondre en fonction du sujet du questionnaire. Des considérations sociales entrent donc en compte dans la décision des individus de répondre aux questionnaires.



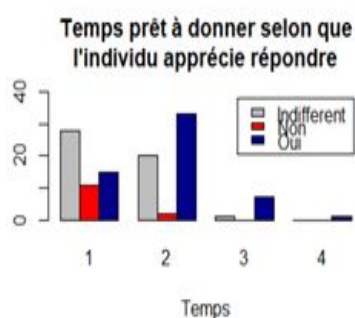
Cet histogramme nous donne les réponses de notre échantillon à la question « à quelle(s) personne(s) répondez-vous ? » dans le cas où l'individu répond au questionnaire en fonction de la personne qui lui soumet. On constate que la plupart des individus répondent d'abord à leur cercle « proche » qui comprend leur famille, leurs amis et leurs collègues de travail. Les institutions telles que l'INSEE et les ONG passent au second plan, tandis que très peu d'individus répondent aux entreprises privées dans notre échantillon.

De plus, une rencontre physique avec l'interrogateur (par exemple dans le cadre d'un questionnaire sur la visite d'un lieu culturel) peut également impacter la motivation de répondre des individus. Par exemple dans notre échantillon, 46% des individus se sentent plus obligés de répondre quand ils rencontrent physiquement l'interrogateur.

Le profil des personnes qui aiment répondre à un questionnaire

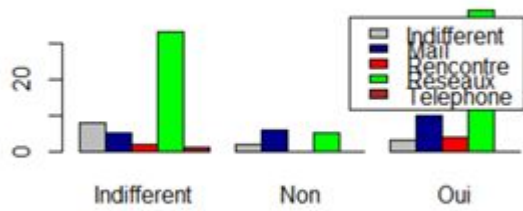
Sur l'ensemble de notre échantillon, 48% déclarent aimer répondre à un questionnaire. Seuls 11% n'aiment pas y répondre et 41% sont indifférents.

Pourtant sur l'ensemble de notre échantillon, 83 % déclarent avoir déjà arrêté de répondre à un questionnaire. On peut donc avancer que certaines caractéristiques des questionnaires peuvent décourager un individu, qui aime pourtant y répondre, de finaliser sa réponse.



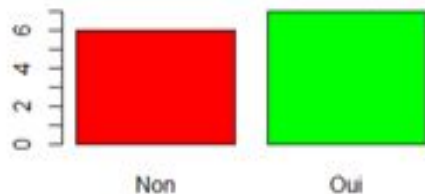
De plus, 90% des individus sont prêts à accorder moins de 10 minutes pour répondre à un questionnaire. Comme attendu, ceux qui aiment répondre à un questionnaire sont prêts à donner en moyenne plus de temps pour répondre à un questionnaire que ceux qui sont indifférents ou qui n'aiment pas répondre. Mais parmi ceux qui aiment répondre à un questionnaire très peu sont prêts à donner plus de 10 minutes pour y répondre. Donc même pour les individus qui aiment répondre à un questionnaire, ceux-ci ne veulent y passer que relativement peu de temps. (1: "Moins de 5 min", 2: "5 à 10 min", 3: "10 à 15 min", 4: "+ de 15 min")

Réseaux préférés selon l'appréciation de répondre à un questionnaire



Appréciation répondre

Influence d'une récompense sur les individus n'aimant pas répondre



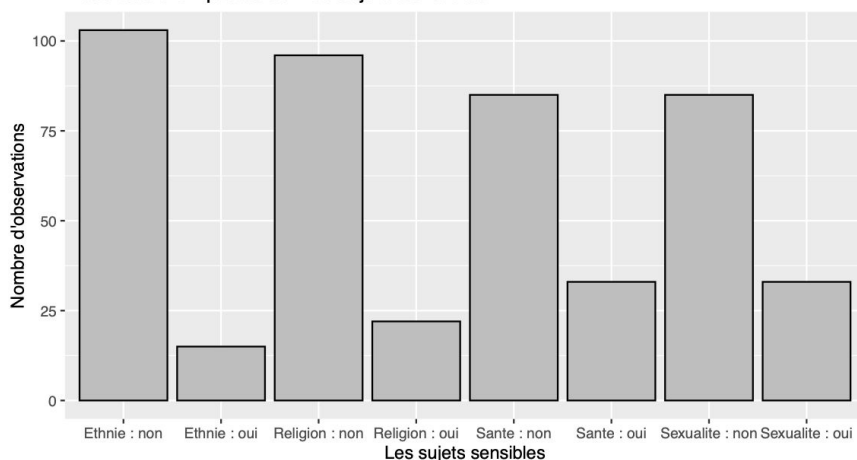
Parmi ceux qui sont indifférents ou qui aiment répondre à un questionnaire, les individus préfèrent en grande majorité être interrogé via les réseaux sociaux. Cette observation est à mettre en lien avec le fait que les individus interrogés sont relativement jeunes. La majorité des individus indifférents face à la façon de se faire soumettre un questionnaire sont aussi indifférents au fait de répondre à un questionnaire.

Parmi ceux qui n'aiment pas répondre aux questionnaires plus de la moitié verrait leur volonté de répondre à un questionnaire affectée s'ils se voyaient offrir une récompense comme un bon de réduction. La promesse d'une récompense semble plutôt efficace pour augmenter le nombre de répondants.

Les sujets sensibles

La commission nationale de l'informatique et des libertés (CNIL) stipule : "il est interdit de collecter ou de traiter des données à caractère personnel qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci." Nous n'avons pas dérogé à la règle, mais nous nous sommes intéressés à l'opinion de nos enquêtés sur ces thématiques, qu'on nommera dans cette partie "les sujets sensibles" (Religion, Ethnie, Santé, Sexualité).

Résultats d'enquêtes sur les sujets sensibles



La tendance dégagée par notre enquête est que la population ne se sentirait pas dérangée par une question concernant un sujet sensible au sens de la CNIL : On a une moyenne de 78% de "non" pour les 4 sujets sensibles (graphique ci contre). Contre une moyenne de 22% de "oui" pour les 4 sujets sensibles.

sentiraient pas dérangé par une question sur leur ethnie. La barre à sa droite indique que 15 personnes ont répondu qu'elles se sentiraient dérangé par une question sur leur ethnie).

(Note de lecture graphique : La barre la plus à gauche indique que 103 personnes ont répondu qu'elles ne se

Si on creuse un peu : 57% des enquêtés ont répondu un **oui global** (1). On ne note pas de différence(s) de réponse selon le sexe : proportionnellement à la taille de l'échantillon, 58% des hommes ont répondu un non global contre 56% des femmes. Davantage sur l'âge : 35% chez les 26 ans et plus contre 60% chez les 25 ans et moins (on pourrait penser à un effet de mœurs, mais nous n'avons pas les résultats suffisant pour conclure, en termes de questions posées ou de taille d'échantillon). A l'opposé, le **non global** (2) concernent 8,4% de l'échantillon, sans différences notables selon le sexe ou le groupe d'âge.



Le sexe et le groupe d'âge ne peuvent pas expliquer les ressentis face aux sujets sensibles. Néanmoins notre population ne montre pas de blocage face à des sujets sensibles au sens de la CNIL.

(1) Oui global : avoir répondu « oui » à chacun des 4 thèmes étudiés. (2) Non global : avoir répondu « non » à chacun des 4 thèmes étudiés.

Note de lecture du graphique : La première barre montre que 67 personnes ont répondu un "non global", parmi eux, 22 sont des hommes.

On peut élargir l'argument : la CNIL a interdit la collecte d'informations pour des sujets sensibles. Mais existe-t-il d'autres sujets que la population juge intrusifs ?

Parmi les 83% de notre échantillon ayant déjà arrêté de répondre à un questionnaire en cours de route, 25% de ces 83% ont arrêté en raison de question(s) jugée(s) intrusive(s). Ceci montre que d'autres sujets que ceux interdits de collecte par la CNIL peuvent déplaire. Et paradoxalement, parmi ces 25%, 1 personne sur 3 avait répondu un "non global" aux sujets sensibles, contre 1 une personne sur 12 ayant répondu un "oui global".

Conclusion

Nous pouvons retenir de notre étude que 89% des individus aiment répondre à un questionnaire ou y sont indifférents. La majorité de ces individus préfèrent être soumis à des questionnaires à travers les réseaux sociaux et les mails. Pour cause on remarque une préférence pour les questionnaires courts et faciles à répondre. Nous avons aussi pu dégager un fort impact du sujet et de l'interrogateur sur l'envie de répondre des individus. Enfin, nous nous sommes intéressés aux questions sensibles interdits d'inclure dans un questionnaire et nous avons pu observer qu'une majorité n'est pas réticente au fait d'y répondre.

Cependant notre étude a des limites intrinsèquement liées à notre échantillon d'analyse. En effet, notre échantillon est de taille limitée (118 individus). De plus, les femmes y sont surreprésentées ainsi que les personnes de 18 à 25 ans et donc les étudiants. Cela a pu biaiser principalement nos résultats concernant les avis aux questions intrusives ainsi que les pratiques d'utilisation des différents moyens de diffusion des questionnaires. Nous avons aussi été limité dans l'analyse de possibles divergences selon l'âge des individus, leur sexe ou encore leur CSP.

Annexe 1 : Le code R

```
library(tidyr)
library(stringr)
library(dplyr)
library(ggplot2)
library(ggformula)

#Graphique des préférences

pref <- theme_minimal()+theme(panel.grid=element_blank(), axis.title.x = element_blank(),
axis.title.y = element_blank(), axis.text.x = element_blank(), axis.text.y = element_blank())

Q %>% gf_bar(~1, fill=~Preferences,xlab="",ylab="") + gf_theme(pref) + coord_polar('y', start = 0)

#Recodage données âge

Q$Age <- dplyr::recode(Q$Age, '0-18'="25-", '18-25' = "25-", '26-35' = "25+", '36+' = "25+")

#Statistiques sur l'échantillon

Q %>% count(Sujet)
Q %>% count(Interrogateur)

#Graphique des interrogateurs préférés

bar <- Q %>% select(Ami, Famille, Colleague, Institutions, Entreprises, ONG) %>% na.omit()

step <- sapply(bar, is.logical)
bar[,step] <- lapply(bar[,step], as.numeric)

bar$Ami <- dplyr::recode(bar$Ami, "1" = "Amis")
bar$Famille <- dplyr::recode(bar$Famille, "1" = "Famille")
bar$Colleague <- dplyr::recode(bar$Colleague, "1" = "Collegues")
bar$Institutions <- dplyr::recode(bar$Institutions, "1" = "Institutions")
bar$Entreprises <- dplyr::recode(bar$Entreprises, "1" = "Entreprises privees")
bar$ONG <- dplyr::recode(bar$ONG, "1" = "ONG")
```

```

diag <- bar %>% gather("key","value", Ami, Famille, Colleague, Institutions, Entreprises, ONG)

diag %>% na.omit() %>% gf_bar(~value)

#Statistiques sur l'échantillon

int <- Q %>% select(Sujet, Arret, Nointeret)%>% filter(Sujet=="Oui") %>% filter(Arret=="Oui")
int %>% count(Nointeret)

ren <- Q %>% select(Age, Sexe, Rencontre) %>% filter (Rencontre=="Oui")
ren %>% count(Age)
ren%>% count(Sexe)

Q %>% count(Rencontre)

#Rerecodage données âge

Q$Age <- dplyr::recode(Q$Age, '0-18'="0-18", '18-25' = "18-25", '26-35' = "26-35", '36+' = "36+")

#Parmi ceux qui *aiment* répondre à un questionnaire, le temps prêt
#à donner à un questionnaire est * majoritairement compris entre 5 et 10 minutes*.

Q$Temps <- dplyr::recode(Q$Temps, '0-5'="1", '5-10' = "2", '10-15' = "3", '15-' = "4")
temps <- table( Q$Appreciation, Q$Temps)
barplot(temps, main=" Temps prêt à donner selon que
l'individu apprécie répondre",
        xlab="Temps",
        col=c("Grey" ,"red", "darkblue"),
        legend = rownames(temps), beside=TRUE, ylim =c(0,40))

#aimer répondre et réseaux préférés.
tmp <- table(Q$Preferences, Q$Appreciation)

barplot(tmp, main="Réseaux préférés selon l'appréciation
de répondre à un questionnaire",
        xlab="Apprécier répondre", col=c("Grey", "darkblue","red", "Green", "Brown"),
        legend = rownames(tmp), beside=TRUE)

#Quel "genre" aime les questionnaires ?
questio_genre <- xtabs(~Sexe, data=Q, subset = Appreciation == "Oui")

barplot(questio_genre,col=c("Red", "Blue"), main = "Aimer répondre à un questionnaire
selon le sexe")

```

#Influence d'une récompense sur les individus n'aimant pas répondre à un questionnaire.

```
rec <- xtabs(~Recompense, data = Q, subset = Appreciation == "Non")  
barplot(rec, col=c("Red", "Green"), main = "Influence d'une récompense sur les  
individus n'aimant pas répondre")
```

#Graphiques des résultats

```
a <- Q %>%  
  select(Ethnie,Religion,Sexualite,Sante, Sexe)  
  
a$Ethnie <- dplyr::recode(a$Ethnie, "Oui" = "Ethnie : oui", "Non"="Ethnie : non")  
a$Religion <- dplyr::recode(a$Religion, "Oui" = "Religion : oui", "Non"="Religion : non")  
a$Sexualite <- dplyr::recode(a$Sexualite, "Oui" = "Sexualite : oui", "Non"="Sexualite : non")  
a$Sante <- dplyr::recode(a$Sante, "Oui" = "Sante : oui", "Non"="Sante : non")  
a$Sexe <- dplyr::recode(a$Sexe, "H" = "H", "F"="F")
```

```
aa<- a %>% gather("key","value",Religion,Sexualite,Sante,Ethnie)  
aa %>%  
  gf_bar(~value, color = "Black", fill = "Grey", size = 0.5, xlab = "Les sujets sensibles",  
    ylab = "Nombre d'observations",  
    title = "Résultats d'enquêtes sur les sujets sensibles")
```

#Moyenne des "non"

```
Q %>% filter(Sexualite%in%c("Non")) #85 non  
Q %>% filter(Sante%in%c("Non"))#85 non  
Q %>% filter(Religion%in%c("Non")) #96 non  
Q %>% filter(Ethnie%in%c("Non"))#103 non  
somme <-c(85,85,96,103)  
moyenne <- mean(somme)  
pourcentage <- moyenne/118 #78% de non en moyenne.
```

#Moyenne des "oui"

```
Q %>% filter(Sexualite%in%c("Oui")) # 33 "oui"  
Q %>% filter(Sante%in%c("Oui"))# 33 "oui"  
Q %>% filter(Religion%in%c("Oui")) #22 "oui"  
Q %>% filter(Ethnie%in%c("Oui"))#15 "oui"  
oui<-c(33,33,22,15)  
moy<-mean(oui)  
enpour100 <- moy/118 #22% de oui en moyenne.
```

#construction du "non global":

```
Q %>%  
  select(Ethnie,Religion,Sexualite,Sante, Sexe,Age) %>%  
  filter(Ethnie%in%c("Non")) %>%  
  filter(Religion%in%c("Non")) %>%
```

```

filter(Sexualite%in%c("Non")) %>%
filter(Sante %in%c("Non")) %>% #67 personnes ont répondu "non" aux 4 questions
filter(Sexe%in%c("H"))# parmi ces 67, 22 sont des hommes
#+nombre selon l'âge.
Q %>% filter(Age%in%c("0-18","18-25"))#60 réponses de moins de 25 ans.

```

#Construction du "oui global"

```

Q %>%
select(Ethnie,Religion,Sexualite,Sante, Sexe,Age) %>%
filter(Ethnie%in%c("Oui")) %>%
filter(Religion%in%c("Oui")) %>%
filter(Sexualite%in%c("Oui")) %>%
filter(Sante %in%c("Oui")) %>% #10 personnes ont répondu "oui" aux 4 questions
filter(Sexe%in%c("H")) %>%# parmi ces 10, 5 sont des hommes.
#+nombre selon l'âge.
filter(Age%in%c("0-18","18-25"))#6 personnes de moins de 25.

```

#Graphique oui/non global selon le sexe.

#On construit un dataframe à partir des résultats des oui et non global.

```

global<- data.frame(Réponse =
c("Oui","Oui","Oui","Oui","Oui","Oui","Oui","Oui","Oui","Oui","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non",
"Non","Non","Non","Non",
"Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non",
"Non","Non","Non","Non",
"Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non","Non",
"Non"),
Sexe = c("F","F","F","F","F","H","H","H","H","H",
"F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F",
"F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F","F",
"H","H","H","H","H","H","H","H","H","H","H","H","H","H","H","H","H","H","H","H","H","H"))
global %>% count(Réponse, Sexe) %>%
mutate(Réponse = factor(Réponse)) %>%
ggplot(aes(Réponse, n)) +
geom_col(aes(fill = Sexe))

```


#Ouverture aux questions intrusives.

```
Q %>% select(Intrusive,Arret) %>%  
  filter(Arret%in%c("Oui")) %>% #99 personnes concernés  
  filter(Intrusive%in%c(TRUE))#24 personnes concernés  
nb_arrêt<- 99/118  
nb_arrêt_intrusive<- 24/99  
nb_arrêt_intrusive_total<- 24/118
```

#Or

```
Q %>%select(Intrusive,Arret,Ethnie,Religion,Sexualite,Sante) %>%  
  filter(Arret%in%c("Oui")) %>%  
  filter(Intrusive%in%c(TRUE)) %>%  
  filter(Ethnie%in%c("Non")) %>%  
  filter(Religion%in%c("Non")) %>%  
  filter(Sexualite%in%c("Non")) %>%  
  filter(Sante%in%c("Non"))#8 personnes concernés.
```