

Questionnaire étudiant

Sebri, JcB

19/02/2015

Contents

1	Questionnaire étudiant	2
1.1	Etablissements participant:	2
1.2	Age	2
1.3	Sexe	4
1.4	Age et sexe	5
1.5	Q1- Pour ce cours, vous avez pris des notes	7
1.6	Q2- Pendant ce cours, vous avez complété la prise de notes par (plusieurs réponses possibles)	7
1.7	Q3- Quels sont les outils numériques que vous aviez avec vous pendant ce cours? (plusieurs réponses possibles)	7
1.8	Q4- Pendant ce cours (en dehors des temps de pause éventuels), vous avez utilisé votre téléphone pour (plusieurs réponses possibles):	15
1.9	Q5- A quelle fréquence, avez-vous utilisé votre téléphone PENDANT ce cours (en dehors des temps de pause éventuels) pour prendre des notes ou chercher sur internet des informations au sujet du cours ?	16
1.10	Q6- A quelle fréquence, avez-vous utilisé votre téléphone PENDANT ce cours (en dehors des temps de pause éventuels) pour faire autre chose que prendre des notes ou chercher sur internet des informations au sujet du cours?	17
1.11	Q7- Pendant ce cours (en dehors des temps de pause éventuels), vous avez utilisé votre tablette et/ ou votre ordinateur pour (plusieurs réponses possibles):	18
1.12	Q8- A quelle fréquence, avez-vous utilisé votre tablette, et/ ou votre ordinateur PENDANT ce cours (en dehors des temps de pause éventuels) pour prendre des notes ou chercher sur internet des informations au sujet du cours ?	19
1.13	Q9- A quelle fréquence, avez-vous utilisé votre tablette, et/ ou votre ordinateur PENDANT ce cours (en dehors des temps de pause éventuels) pour faire autre chose que prendre des notes ou chercher sur internet des informations au sujet du cours ?	20
2	Questions supplémentaires	21
3	Résultats selon la promotion	21
3.1	Question Q4	21
3.2	Question Q7	27
4	Information de session	32

Version du: **Sun Jun 14 19:44:25 2015**

1 Questionnaire étudiant

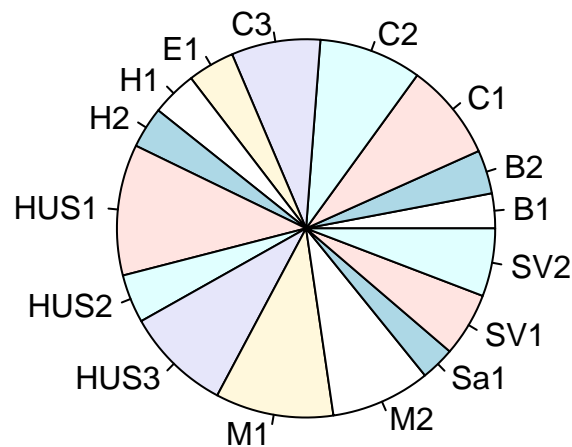
```
[1] "Etab"      "Etud"      "Q1"        "Q2.1"      "Q2.2"      "Q2.3"      "Q2.4"
[8] "Q2.5"      "Q2.6"      "Q2.7"      "Q3.1tpc"   "Q3.2sp"    "Q3.3tab"   "Q3.4ord"
[15] "Q4.1"      "Q4.2"      "Q4.3"      "Q4.4"      "Q4.5"      "Q4.6"      "Q4.7"
[22] "Q4.8"      "Q4.9"      "Q4.10"     "Q4.11"     "Q4.12"     "Q4.13"     "Q4.14"
[29] "Q4.15"     "Q4.16"     "Q5"        "Q6"        "Q7.1"      "Q7.2"      "Q7.3"
[36] "Q7.4"      "Q7.5"      "Q7.6"      "Q7.7"      "Q7.8"      "Q7.9"      "Q7.10"
[43] "Q7.11"     "Q7.12"     "Q7.13"     "Q7.14"     "Q7.15"     "Q7.16"     "Q8"
[50] "Q9"        "Q10"       "Q11"
```

Le fichier comporte:

- 1446 lignes
- 52 variables

1.1 Etablissements participant:

```
B1  B2  C1  C2  C3  E1  H1  H2  HUS1 HUS2 HUS3  M1  M2  Sa1  SV1
43  54  120 127 110 58  56  51  162  60  131 146 123  42  79
SV2
84
```



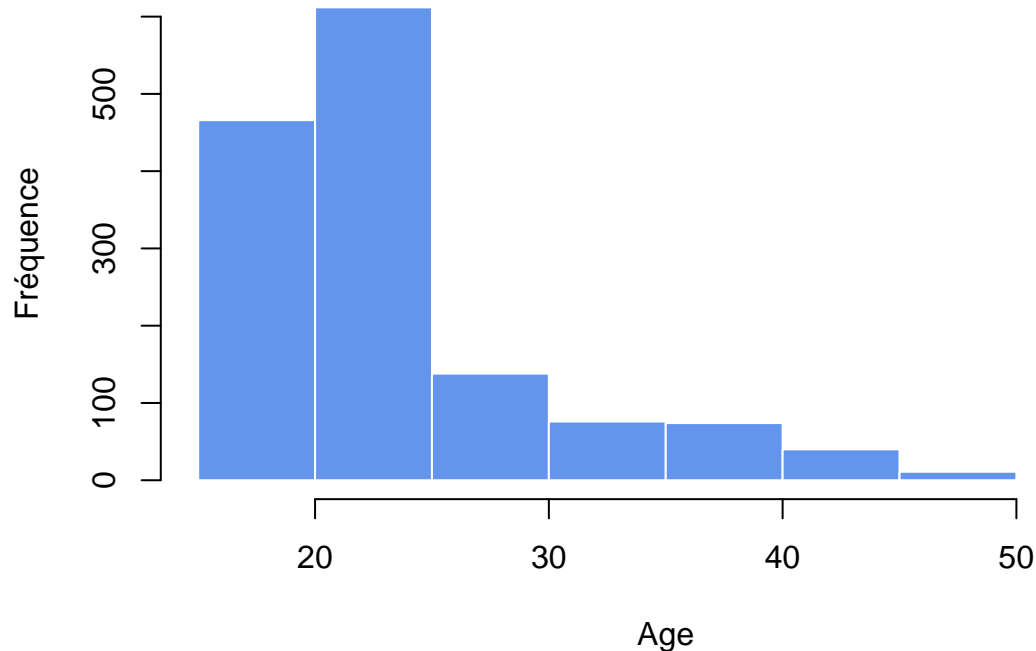
1.2 Age

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
17.0	20.0	22.0	24.1	25.0	53.0	27

1.2.1 Générations

```
# génération
# Z = 15 à 20 ans
# Y = 20 à 35 ans
```

Histogramme de l'age



```
# X > 35 ans
```

```
age <- c(15, 20, 35, 60)
g <- cut(d1$Q11, age)
summary(g)
```

```
## (15,20] (20,35] (35,60] NA's
##      466      826      127      27
```

```
g2 <- cut(d1$Q11, age, labels = c("Z", "Y", "X"))
summary(g2)
```

```
##      Z      Y      X NA's
##  466  826  127   27
```

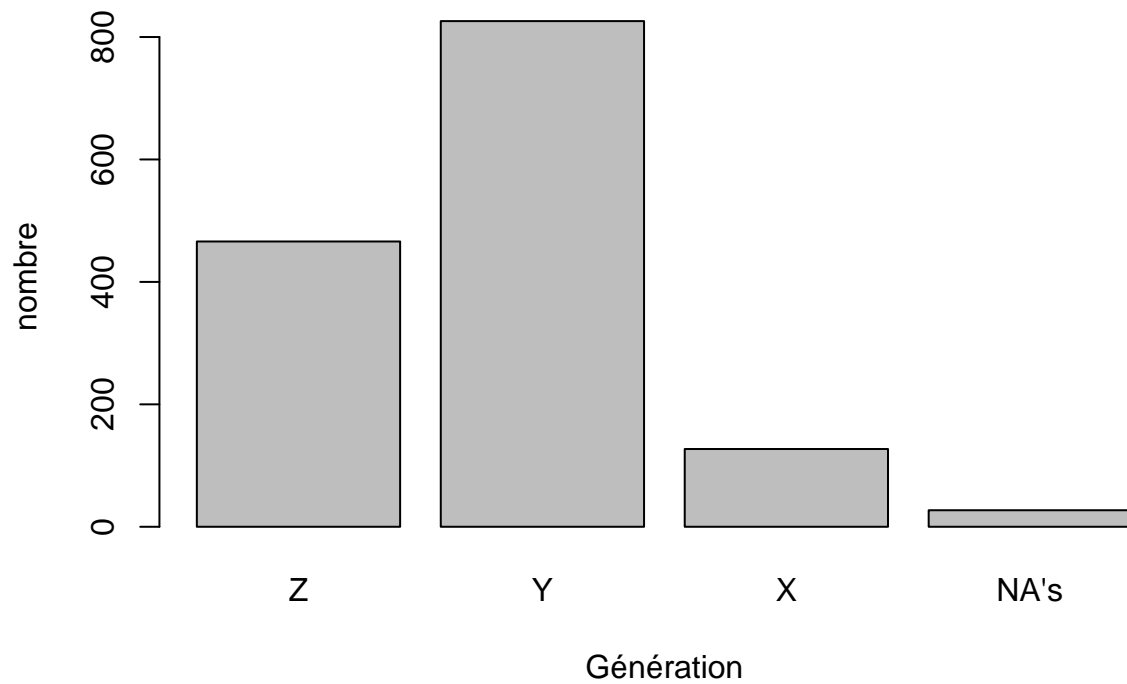
```
# ajout d'une colonne GENERATION
```

```
d1$GENERATION <- g2
factor2table(d1$GENERATION)
```

```
##           Z           Y           X  NA's
## nombre    466.00  826.00  127.00  27.00
## proportion  32.23  57.12   8.78   1.87
```

```
barplot(summary(d1$GENERATION), xlab = "Génération", ylab = "nombre", main = "Répartition des générations")
```

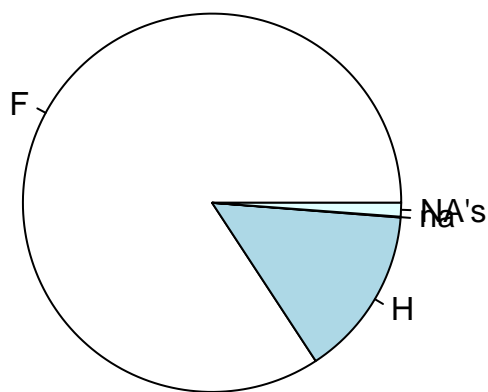
Répartition des générations au sein des étudiants



1.3 Sexe

F	H	na	NA's
1218	210	1	17

Sexe



Test de la routine **factor2table**

```
f <- factor2table(d1$Q10, digit=2, col=c("femmes", "hommes", "inconnu"))
f
```

```
##          femmes hommes inconnu
## nombre  1218.00 210.00   18.00
## proportion  84.23  14.52    1.24
```

Table 1: Sexe des participants

	femmes	hommes	inconnu
nombre	1218.00	210.00	18.00
proportion	84.23	14.52	1.24

Sous forme de tableau avec **kable**:

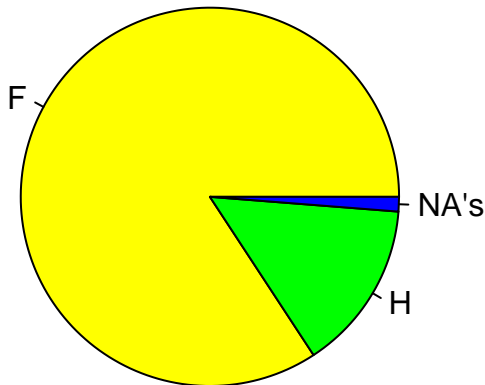
Sous forme de tableau avec **xtable**:

% latex table generated in R 3.1.3 by xtable 1.7-4 package % Sun Jun 14 19:44:30 2015

	femmes	hommes	inconnu
nombre	1218.00	210.00	18.00
proportion	84.23	14.52	1.24

Table 2: Sexe des participants

Sexe



Pie chart

1.4 Age et sexe

L'âge des hommes et des femmes sont-ils identiques ? On part de l'hypothèse qu'il n'y a à priori de différence d'âge entre les hommes et les femmes (on appelle cela l'hypothèse nulle ou H_0). Si cette hypothèse est vraie, la différence des moyennes des âges entre les hommes et les femmes devrait être nulle. En pratique cette différence est rarement exactement égale à 0 et le problème est de savoir si le chiffre obtenu est assimilable à 0 ou si au contraire il est trop important pour qu'on puisse se livrer à cette assimilation, auquel cas on est obligé de renoncer à l'hypothèse nulle et accepter l'hypothèse alternative: l'âge des hommes est en moyenne différent de celui des femmes. Pour répondre à la question, on pratique un test statistique pour lequel on définit un écart par rapport à 0. Si le résultat du test tombe dans l'intervalle on admet que la différence de moyenne est assimilable à 0 et on accepte l'hypothèse nulle: pas de différence entre les groupes. Sinon on la rejette. Bien sûr, plus on définit un intervalle important, plus on augmente le risque de se tromper en affirmant qu'il n'y a pas de différence entre les moyennes. C'est ce qu'on appelle le risque de première espèce ou α . Dans les sciences de la santé, ce risque est fixé consensuellement (et arbitrairement) à $5\% = 5/100 = 0.05$ et généralement rapporté sous la forme $p = 0.05$. C'est à dire que j'admet H_0 (pas de différence) en prenant le risque consenti de me tromper dans 5% des cas. En pratique les logiciels calculent la probabilité exacte d'observer par hasard une telle différence entre les deux groupes. Si cette probabilité est supérieure à 0.05 (c'est-à-dire comprise entre 0.05 et 1) on considère que la différence entre les moyennes est un artefact lié au

fluctuation d'échantillonnage et qu'en réalité il n'y a pas de différence entre les groupes. Si au contraire, la probabilité exacte est inférieure à 0.05, on admet qu'elle n'est pas due au hasard et on est obligé d'admettre qu'il y a bien une différence entre les deux groupes. On voit par là le côté arbitraire du petit p, mais il est considéré dans toutes les publications comme un chiffre magique...

Il existe de nombreux tests statistiques. Pour répondre à la question posée, on utilise le test t de Student qui s'applique si:

- on ne compare que 2 groupes (c'est le cas)
- la variable d'intérêt (ici l'âge) suit une loi normale (on va admettre que oui) dans les 2 groupes
- la variance (moyenne des écarts à la moyenne) des 2 groupes est égale (si ce n'est pas le cas, on peut utiliser une variante de test de Student appelée test de Welch).

La colonne sexe (Q10) comporte 3 valeurs: H, F et NR. Il faut éliminer les NR en les transformant en NA pour rendre le test possible

```
d1$Q10 <- toupper(d1$Q10)
d1$Q10[d1$Q10 == "NR"] <- NA
```

Puis faire le test:

```
t <- t.test(d1$Q11 ~ d1$Q10, var.equal = TRUE)
t
```

```
##
## Two Sample t-test
##
## data: d1$Q11 by d1$Q10
## t = -3.7563, df = 1415, p-value = 0.0001795
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -2.7497617 -0.8630462
## sample estimates:
## mean in group F mean in group H
##      23.82645      25.63285
```

```
p.t <- t$p.value
```

On voit que la probabilité exacte d'observer par hasard une telle différence entre les moyennes est égale à 0.0001795. Cette probabilité est très inférieure à 0.05 et donc on rejette l'hypothèse d'égalité des âges. En moyenne, pour cet échantillon, les étudiants hommes sont plus âgés que les étudiantes et cette différence est statistiquement significative.

Comme on peut avoir un doute sérieux sur la normalité de l'âge (voir le graphique des âges ci-dessus), on réalise un test non paramétrique, c'est à dire qui ne fait pas d'hypothèse sur la façon dont la variable est distribuée. Dans le cas particulier on utilise le test de Wilcoxon qui est l'équivalent non paramétrique du test de Student:

```
wilcox.test(d1$Q11 ~ d1$Q10)
```

```
##
## Wilcoxon rank sum test with continuity correction
```

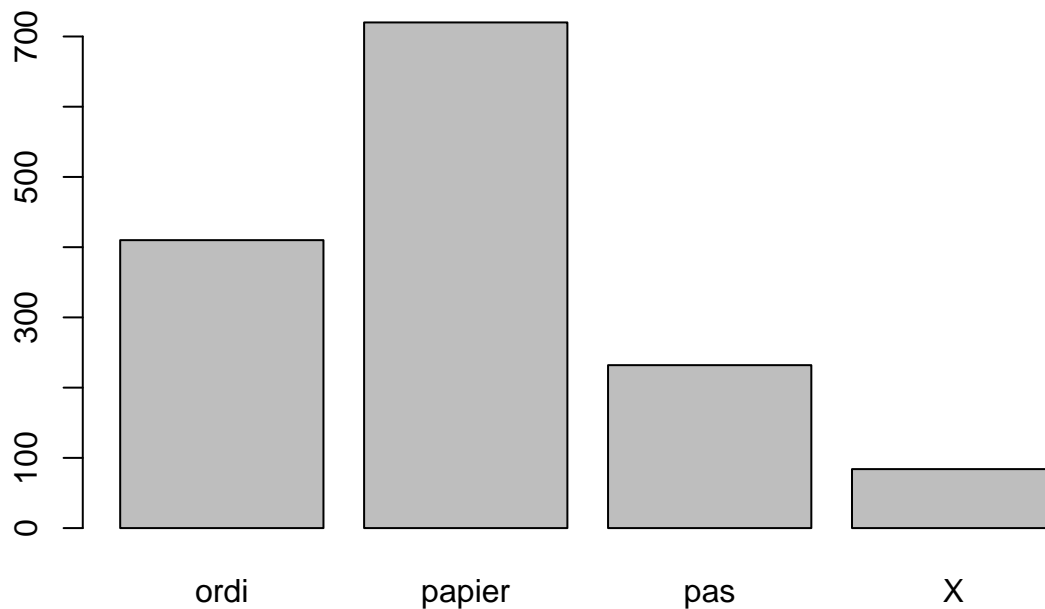
```
##
## data: d1$Q11 by d1$Q10
## W = 97211, p-value = 0.0000002159
## alternative hypothesis: true location shift is not equal to 0
```

On arrive à la même conclusion.

1.5 Q1- Pour ce cours, vous avez pris des notes

	ordi	papier	pas	X
nombre	410.00	720.00	232.00	84.00
proportion	28.35	49.79	16.04	5.81

Support de notes utilisé par l'étudiant



1.6 Q2- Pendant ce cours, vous avez complété la prise de notes par (plusieurs réponses possibles)

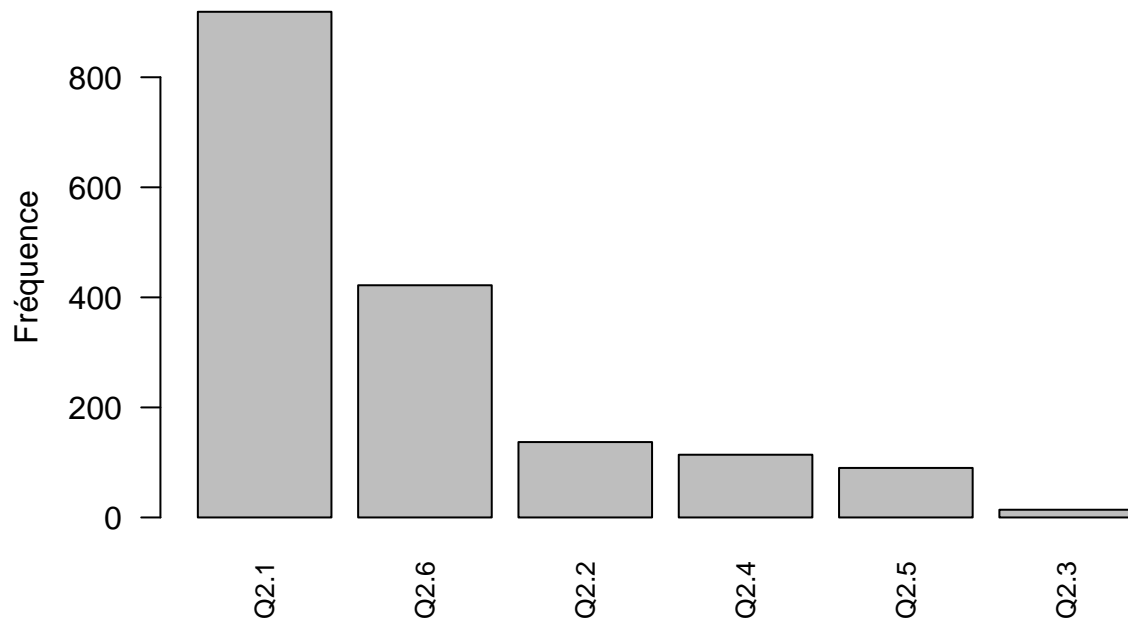
La variable Q2.5 est anormale. Il ne peut y avoir dans la même colonne du texte et des nombres. La colonne ne peut contenir que 1 ou NA. Créer une colonne supplémentaire pour le texte. Par ex. Q2-7.

	0	1	2	3	4
nombre	8.00	1201.00	218.00	17.00	2.00
proportion	0.55	83.06	15.08	1.18	0.14

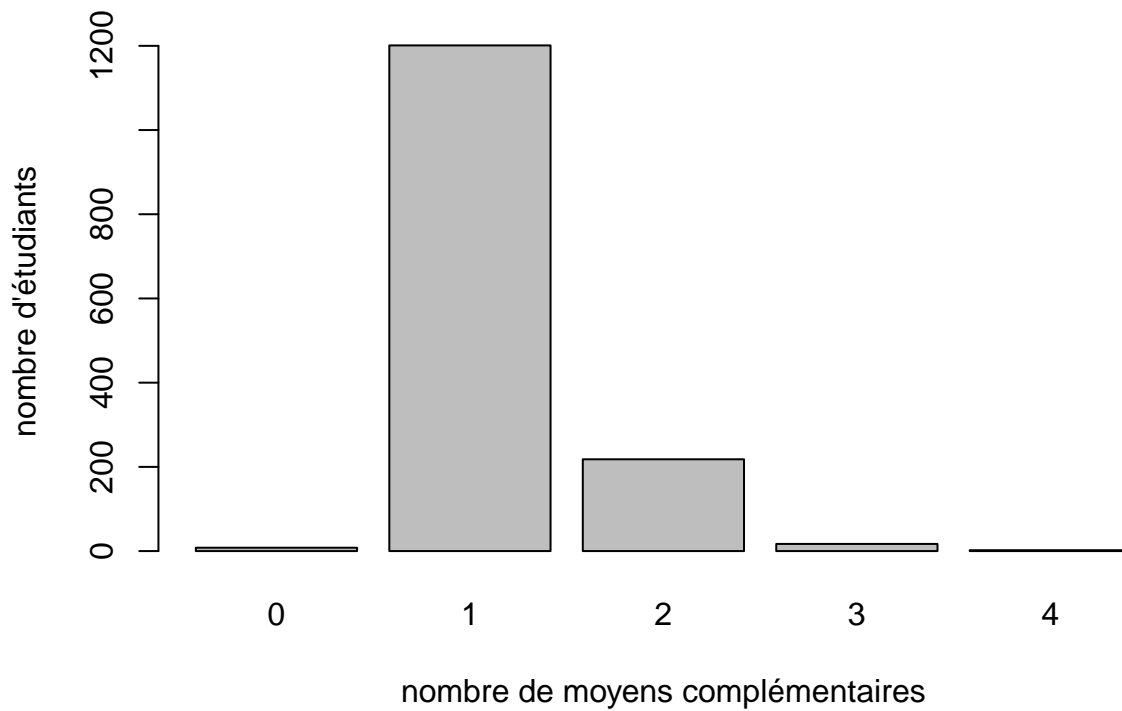
1.7 Q3- Quels sont les outils numériques que vous aviez avec vous pendant ce cours? (plusieurs réponses possibles)

Colonnes 11 à 14

Compléments de la prise de notes



Nombre de compléments aux prises de notes

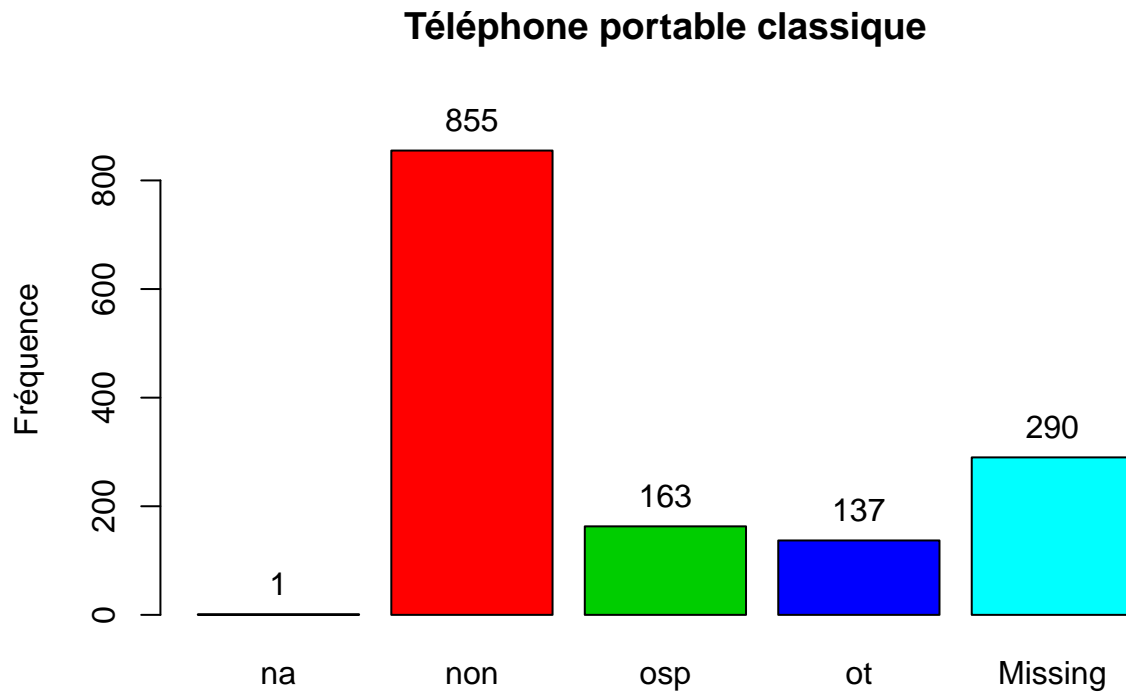


4 types d'outils:

- téléphone portable classique (tpc)
- smartphone (sp)
- tablette (tab)
- ordinateur portable (ord)

ces outils sont-ils disponibles: non, oui, et si oui où:

- sur la table= ot
- dans mon sac ou ma poche= osp

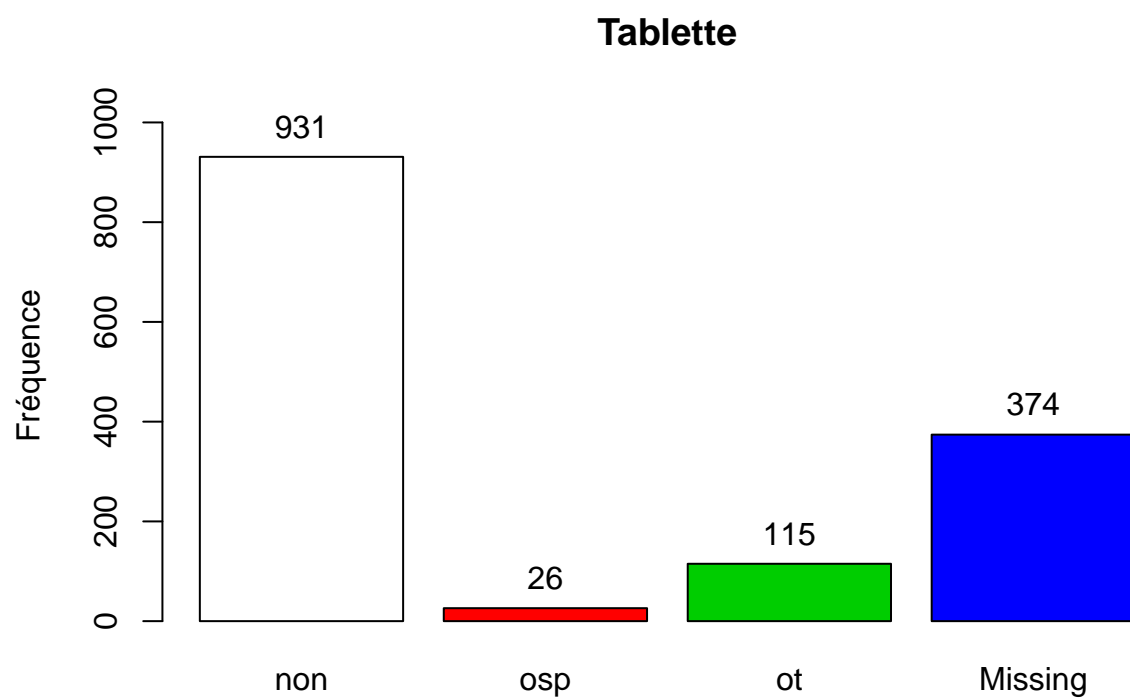
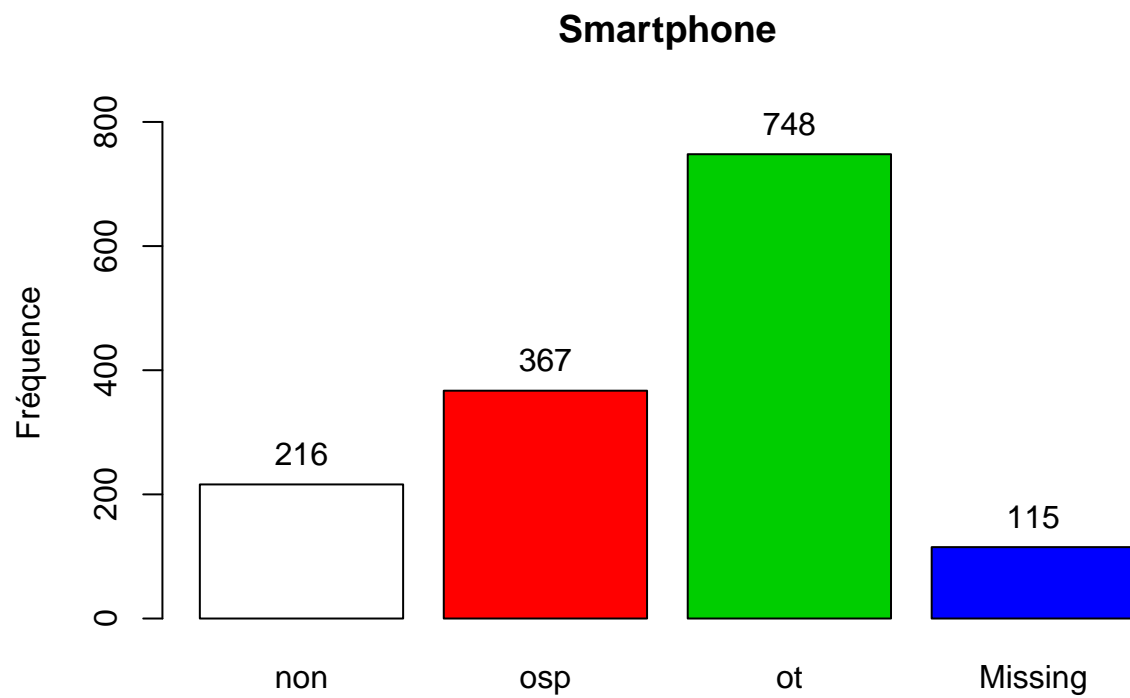


d1\$Q3.1tpc :

	Frequency	%(NA+)	%(NA-)
na	1	0.1	0.1
non	855	59.1	74.0
osp	163	11.3	14.1
ot	137	9.5	11.9
<NA>	290	20.1	0.0
Total	1446	100.0	100.0

d1\$Q3.2sp :

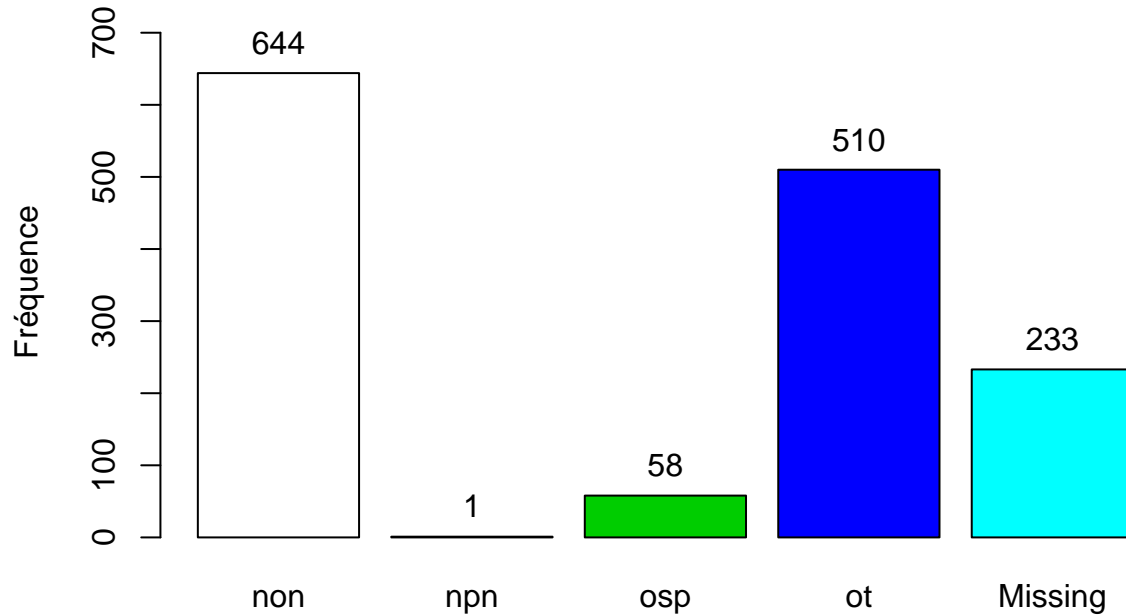
	Frequency	%(NA+)	%(NA-)
non	216	14.9	16.2
osp	367	25.4	27.6
ot	748	51.7	56.2
<NA>	115	8.0	0.0
Total	1446	100.0	100.0



d1\$Q3.3tab :

	Frequency	%(NA+)	%(NA-)
non	931	64.4	86.8
osp	26	1.8	2.4
ot	115	8.0	10.7
<NA>	374	25.9	0.0
Total	1446	100.0	100.0

Ordinateur



d1\$Q3.4ord :

	Frequency	%(NA+)	%(NA-)
non	644	44.5	53.1
nnp	1	0.1	0.1
osp	58	4.0	4.8
ot	510	35.3	42.0
<NA>	233	16.1	0.0
Total	1446	100.0	100.0

Résultats:

- nombre de personnes n'ayant pas répondu à chacune des questions: voir table 3 pp 11

% latex table generated in R 3.1.3 by xtable 1.7-4 package % Sun Jun 14 19:44:39 2015

	Q3.1tpc	Q3.2sp	Q3.3tab	Q3.4ord
nombre	290.00	115.00	374.00	233.00
%	20.06	7.95	25.86	16.11

Table 3: Ne répondent à aucune des 4 questions

- ont un tpc: 300 (20.75 %)

- ont un smartphone: 1115 (77.11 %)
- ont un tpc ET un smartphone: 106 (7.33 %)
- ont un tpc OU un smartphone: 1309 (90.53 %)
- n'ont NI tpc NI sp: 137 (9.47 %)
- ont une tablette: 141 (9.75 %)
- ont un ordinateur portable: 568 (39.28 %)
- ont une tablette ET un ordinateur: 21 (1.45 %)
- ont une tablette OU un ordinateur: 688 (47.58 %)
- n'ont NI ordi, NI tablette: 757 (52.35 %)
- ont un ordinateur ET un smartphone: 475 (32.85 %)
- ne répondent à aucune des 4 questions: 11 (0.76 %)

1.7.1 selon la génération

1.7.1.1 portable classique

```
##
##           Z   Y   X
##   na      0   0   1
##   non 299 495 46
##   osp  36  77 47
##   ot   34  90 11

##   na  non  oui NA's
##     1 855 300 290

##           génération
## possède un tpc   Z   Y   X
##           na      0   0   1
##           non 299 495 46
##           oui  70 167 58

## Warning in chisq.test(table(tpc, d1$GENERATION)): Chi-squared approximation
## may be incorrect

##
## Pearson's Chi-squared test
##
## data:  table(tpc, d1$GENERATION)
## X-squared = 67.0208, df = 4, p-value = 9.651e-14
```

1.7.1.2 smartphone

```
##
##           Z   Y   X
##  non  44 108  57
##  osp 153 186  25
##  ot   254 466  15

##  non  oui NA's
##  216 1115  115

##           génération
## possède un sp   Z   Y   X
##           non  44 108  57
##           oui 407 652  40

##
##  Pearson's Chi-squared test
##
## data:  table(sp, d1$GENERATION)
## X-squared = 147.0316, df = 2, p-value < 2.2e-16
```

1.7.1.3 tablette

```
##
##           Z   Y   X
##  non 315 518  81
##  osp   8  13   4
##  ot   37  76   2

##  non  oui NA's
##  931  141 374

##           génération
## possède un tab   Z   Y   X
##           non 315 518  81
##           oui  45  89   6

##
##  Pearson's Chi-squared test
##
## data:  table(tab, d1$GENERATION)
## X-squared = 4.2748, df = 2, p-value = 0.118
```

Pas de différence entre les génération

1.7.1.4 ordinateur

```

##
##      Z    Y    X
## non 189 374 68
## npn  0   0   1
## osp 15  38   3
## ot  202 278 21

## non oui NA's
## 645 568 233

##      génération
## possède un ord  Z    Y    X
##      non 189 374 69
##      oui 217 316 24

##
## Pearson's Chi-squared test
##
## data:  table(ord, d1$GENERATION)
## X-squared = 23.945, df = 2, p-value = 0.000006316

## [1] "tableau attendu si H0 vraie"

##
## ord      Z      Y      X
## non 215.8049 366.762 49.43314
## oui 190.1951 323.238 43.56686

## [1] "différence observé - attendu"

##
## ord      Z      Y      X
## non -26.804878  7.238015 19.566863
## oui  26.804878 -7.238015 -19.566863

```

1.7.2 création d'une colonne moyens de COMmunication:

Possédez-vous un tpc ou un smartphone:

```

## non oui NA's
## 91 1312 43

```

Selon la génération:

```

##
## non oui
## Z 14 445
## Y 45 756
## X 27 91

```

```
##
##      non   oui
##  Z  3.05 96.95
##  Y  5.62 94.38
##  X 22.88 77.12

##
## Pearson's Chi-squared test
##
## data:  table(d1$GENERATION, d1$COM)
## X-squared = 64.3581, df = 2, p-value = 1.059e-14
```

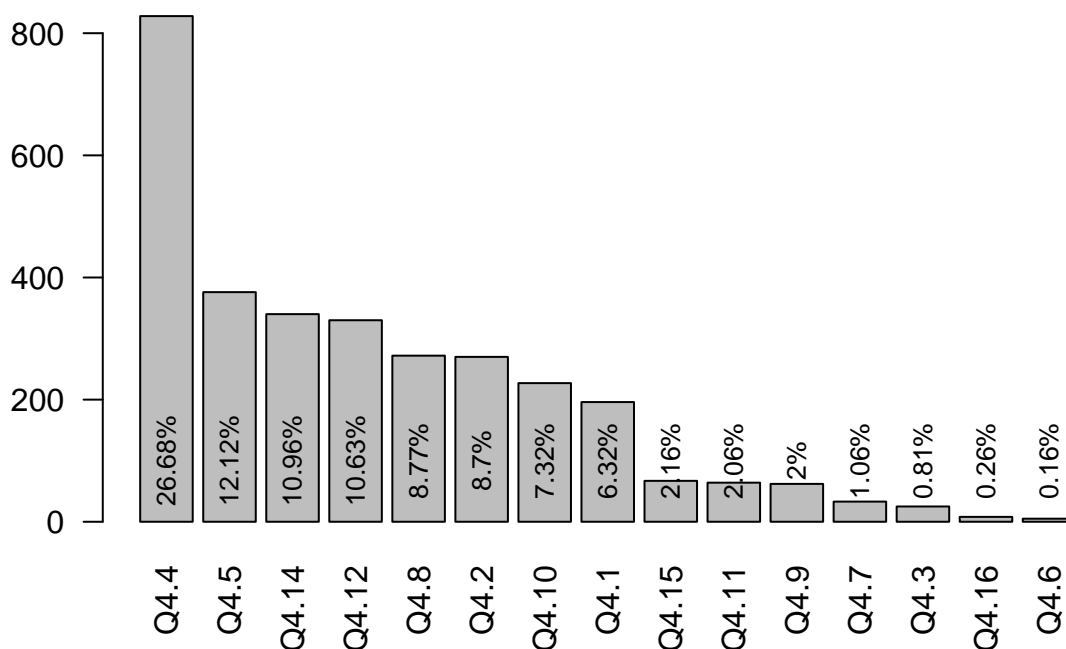
1.8 Q4- Pendant ce cours (en dehors des temps de pause éventuels), vous avez utilisé votre téléphone pour (plusieurs réponses possibles):

question 15 à 30

```
d1 <- read.csv(paste0(path, file1), skip = 1, stringsAsFactors = FALSE)

q4 <- d1[, c(15:26, 28:30)]
q4 <- as.data.frame(sapply(q4, gsub, pattern="NR", replacement="NA"), stringsAsFactors = FALSE)
q4 <- as.data.frame(sapply(q4, as.integer))
a <- apply(q4, 2, sum, na.rm = TRUE)
x <- barplot(sort(a, decreasing = TRUE), las = 2, main = "Utilisation du téléphone pendant le cours")
v <- paste0(sort(round(a*100/sum(a), 2), decreasing = TRUE), "%")
text(x, 100, v, srt=90, cex = 0.8)
```

Utilisation du téléphone pendant le cours

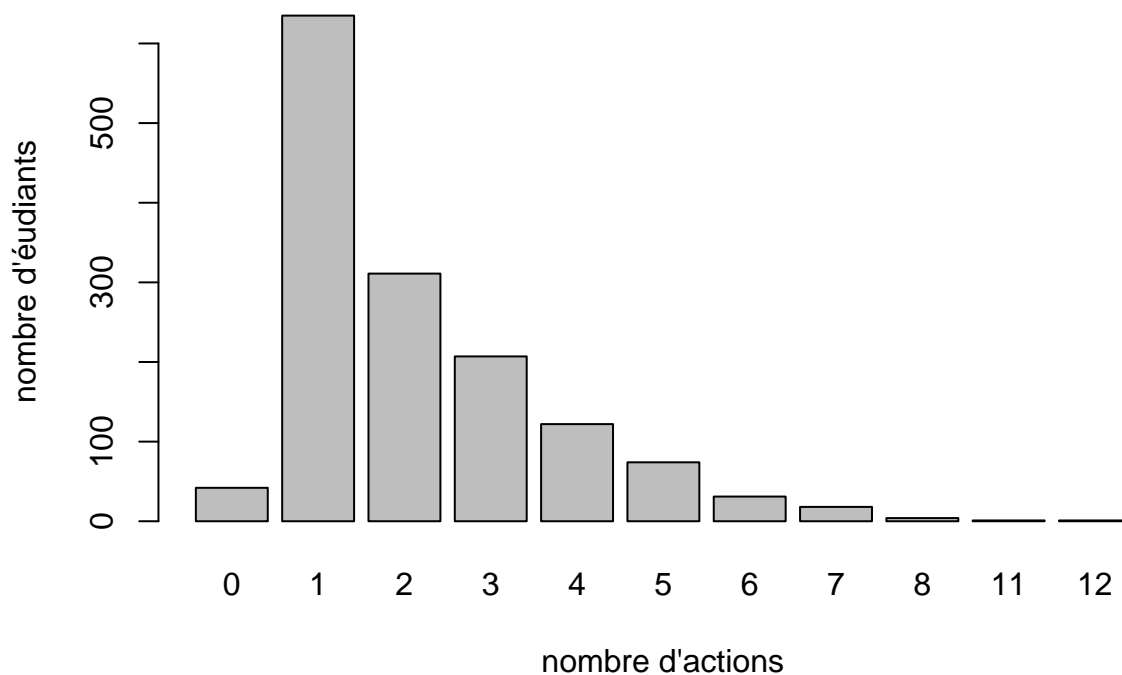


Combien d'actions simultanément:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.000	1.000	2.000	2.146	3.000	12.000

0	1	2	3	4	5	6	7	8	11	12
42	635	311	207	122	74	31	18	4	1	1

nombre d'action pendant le cours



1.9 Q5- A quelle fréquence, avez-vous utilisé votre téléphone PENDANT ce cours (en dehors des temps de pause éventuels) pour prendre des notes ou chercher sur internet des informations au sujet du cours ?

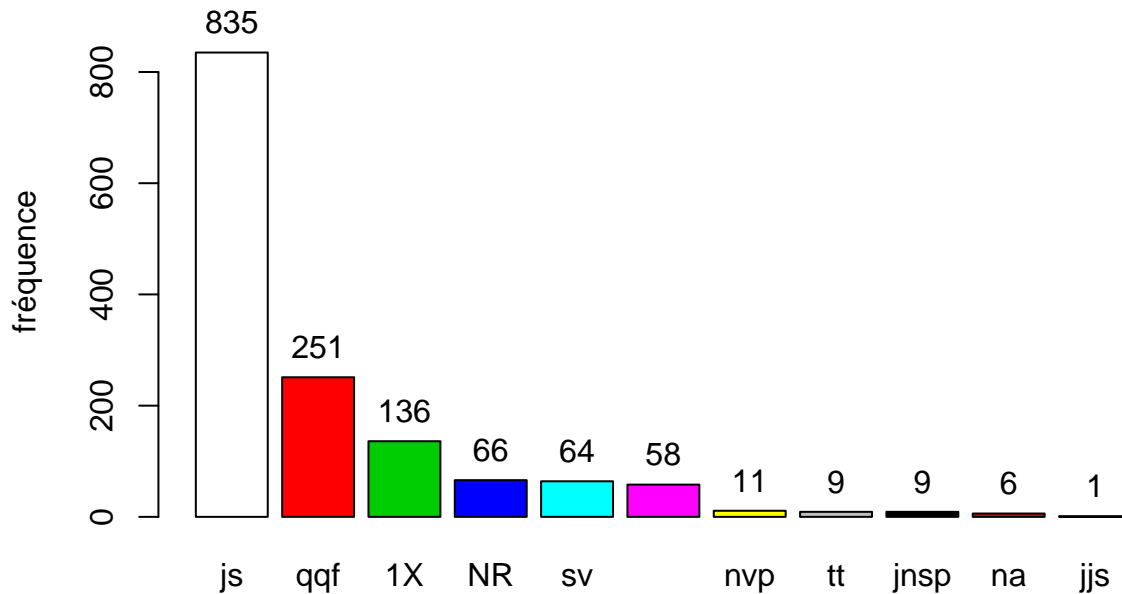
question 31

1X	jjs	jnsp	js	na	NR	nvp	qqf	sv	tt	NA's
136	1	9	835	6	66	11	251	64	9	58

as.factor(d1\$Q5) :

	Frequency	%(NA+)	%(NA-)
js	835	57.7	60.2
qqf	251	17.4	18.1
1X	136	9.4	9.8
NR	66	4.6	4.8
sv	64	4.4	4.6
NA's	58	4.0	0.0
nvp	11	0.8	0.8
jnsp	9	0.6	0.6

Fréquence d'utilisation du téléphone pendant le cours



tt	9	0.6	0.6
na	6	0.4	0.4
jjs	1	0.1	0.1
Total	1446	100.0	100.0

1.10 Q6- A quelle fréquence, avez-vous utilisé votre téléphone PENDANT ce cours (en dehors des temps de pause éventuels) pour faire autre chose que prendre des notes ou chercher sur internet des informations au sujet du cours?

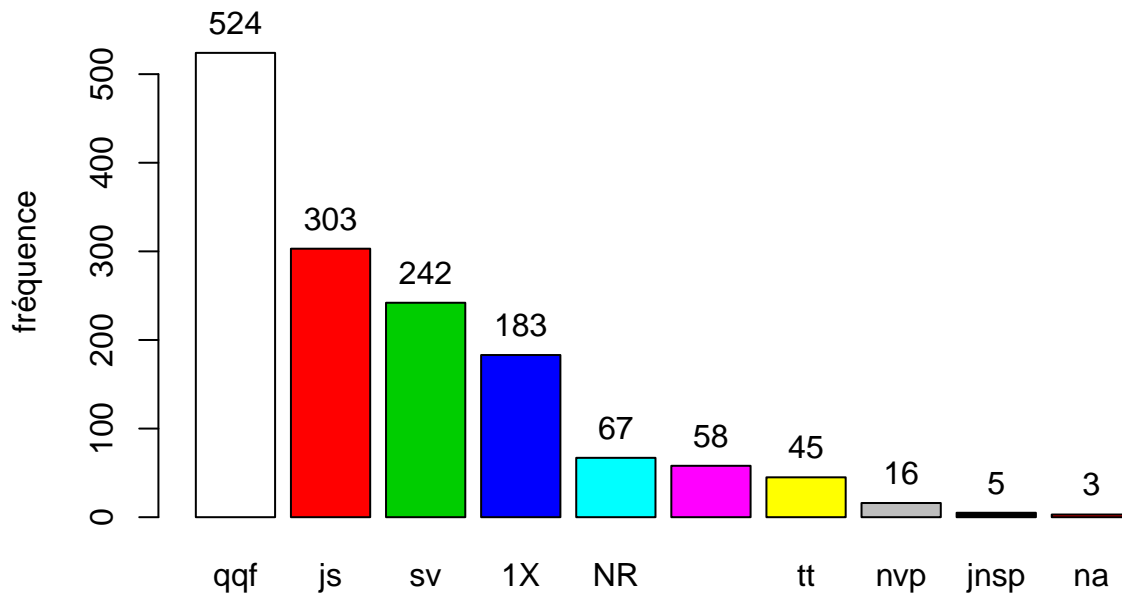
question 32

1X	jnsp	js	na	NR	nvp	qqf	sv	tt	NA's
183	5	303	3	67	16	524	242	45	58

as.factor(d1\$Q6) :

	Frequency	%(NA+)	%(NA-)
qqf	524	36.2	37.8
js	303	21.0	21.8
sv	242	16.7	17.4
1X	183	12.7	13.2
NR	67	4.6	4.8
NA's	58	4.0	0.0
tt	45	3.1	3.2
nvp	16	1.1	1.2
jnsp	5	0.3	0.4
na	3	0.2	0.2
Total	1446	100.0	100.0

Fréquence d'utilisation du téléphone pendant le cours pour faire autre chose



1.11 Q7- Pendant ce cours (en dehors des temps de pause éventuels), vous avez utilisé votre tablette et/ ou votre ordinateur pour (plusieurs réponses possibles):

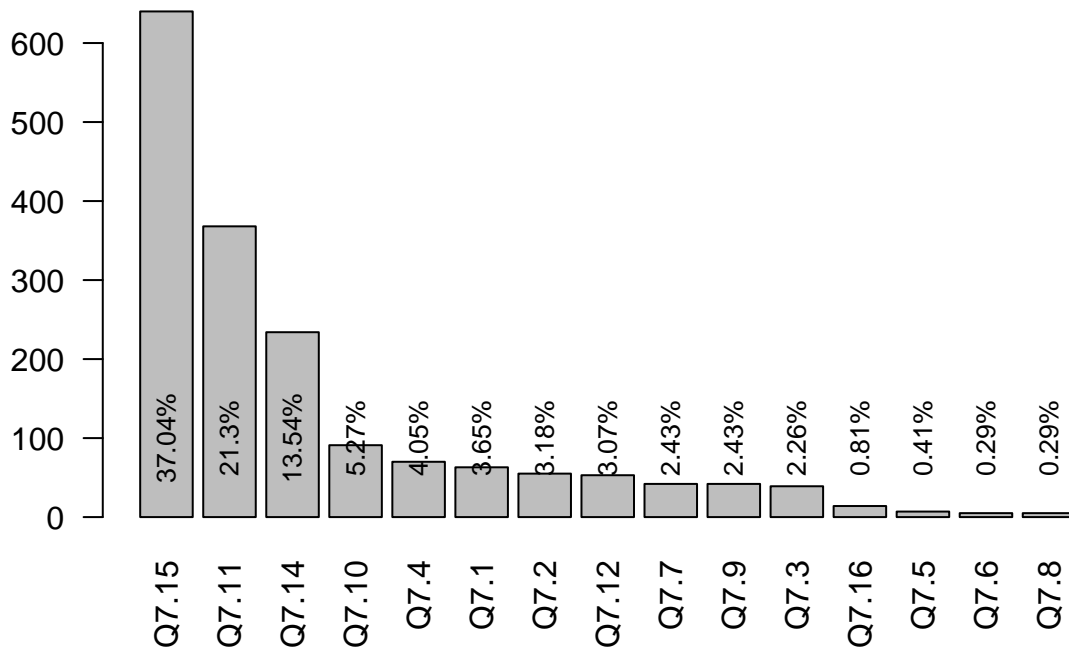
Questions 33 à 48

[1] "Analyse de la colonne Q7.13 (réponse libre)"

	1	cours	dessinerpaint	docmail
1335	1	1	1	1
frfiches	inscr	course	Lemotiv	lire
1	1	1	3	39
orgdossinfo	ppt	pptcours	prepCV	reg_autr_cour
1	1	3	2	1
reg_pptander	reg-heure	reg-photos	reg-ppt	regautre cours
1	1	1	1	1
regcours	regppt	shop	suivre cours	support
1	1	3	1	1
telecharcours	TFE	word	WTFE	NA's
1	3	1	1	36

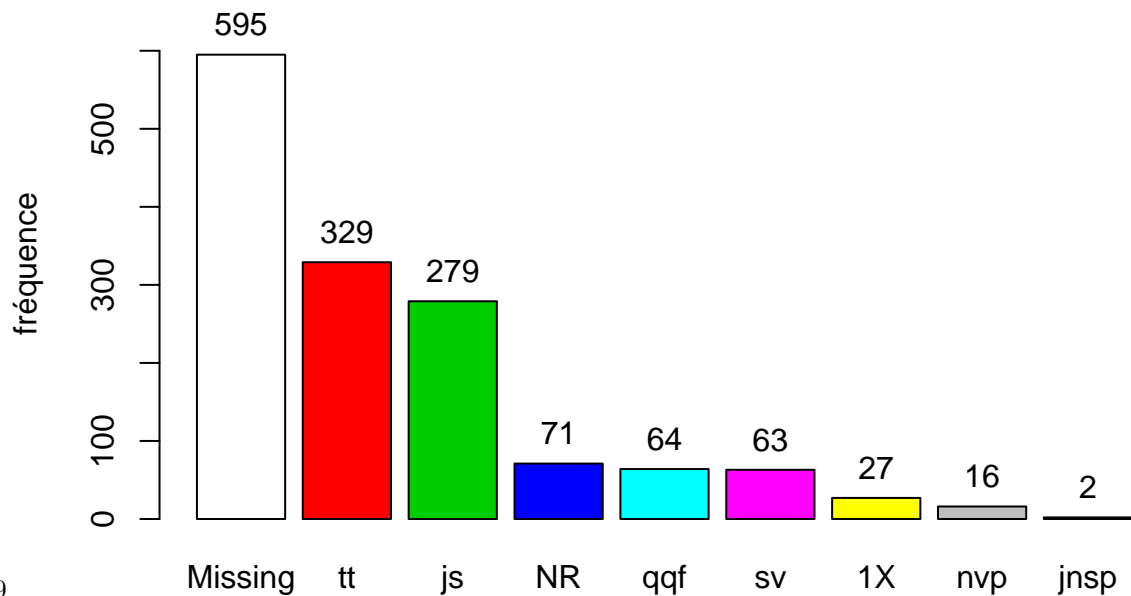
Q7.1	Q7.2	Q7.3	Q7.4	Q7.5	Q7.6	Q7.7	Q7.8	Q7.9	Q7.10	Q7.11	Q7.12
63	55	39	70	7	5	42	5	42	91	368	53
Q7.14	Q7.15	Q7.16									
234	640	14									

Utilisation de la tablette pendant le cours pour:



1.12 Q8- A quelle fréquence, avez-vous utilisé votre tablette, et/ ou votre ordinateur PENDANT ce cours (en dehors des temps de pause éventuels) pour prendre des notes ou chercher sur internet des informations au sujet du cours ?

Fréquence d'utilisation du téléphone pendant le cours pour prendre des notes



question 49

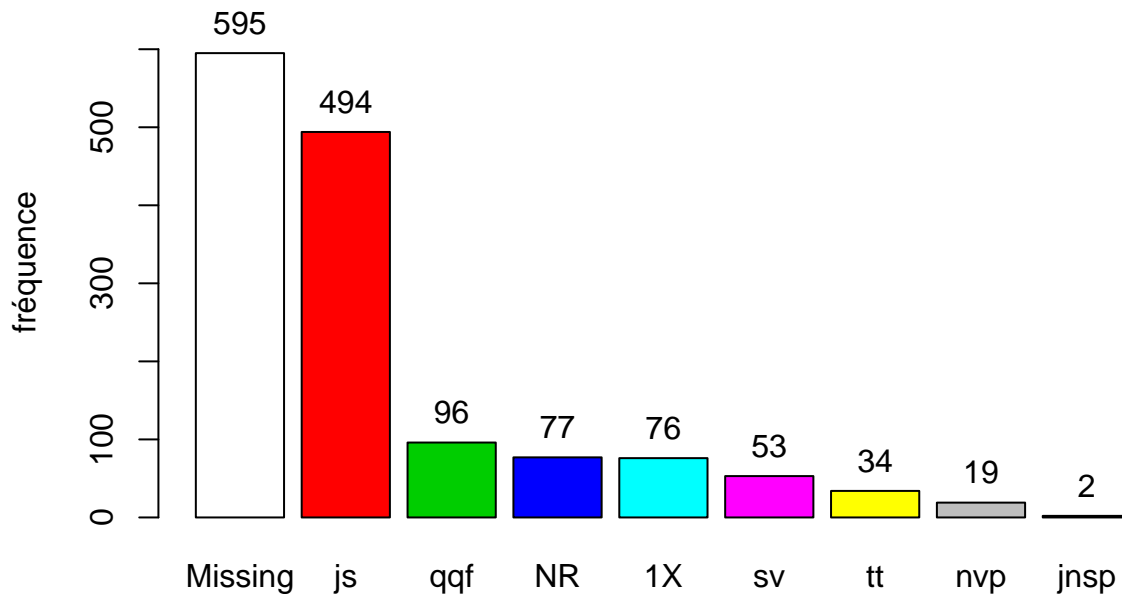
as.factor(d1\$Q8) :

	Frequency	%(NA+)	%(NA-)
NA's	595	41.1	0.0
tt	329	22.8	38.7
js	279	19.3	32.8
NR	71	4.9	8.3
qqf	64	4.4	7.5
sv	63	4.4	7.4
1X	27	1.9	3.2
nvp	16	1.1	1.9
jnsp	2	0.1	0.2
Total	1446	100.0	100.0

1.13 Q9- A quelle fréquence, avez-vous utilisé votre tablette, et/ ou votre ordinateur PENDANT ce cours (en dehors des temps de pause éventuels) pour faire autre chose que prendre des notes ou chercher sur internet des informations au sujet du cours ?

question 50

Fréquence d'utilisation du téléphone pendant le cours pour faire autre chose



```
as.factor(d1$Q9) :
```

	Frequency	%(NA+)	%(NA-)
NA's	595	41.1	0.0
js	494	34.2	58.0
qqf	96	6.6	11.3
NR	77	5.3	9.0
1X	76	5.3	8.9
sv	53	3.7	6.2
tt	34	2.4	4.0
nvp	19	1.3	2.2

jns	2	0.1	0.2
Total	1446	100.0	100.0

2 Questions supplémentaires

- Est-ce que les comportements concernant en particulier la prise de notes étaient différents lorsque les étudiants n'avaient pas de support et savaient qu'ils ne l'auraient jamais à autrement dit, est-ce que l'utilisation des outils numériques est modifiée selon que les étudiants aient ou non accès à un support de cours ? comparaison de la population (HUS 1, E1, C2, C3, B1, B2) avec (Sa1, SV1 SV2, H1 H2, HUS 2 HUS 3, C1,) sur Q4 et Q7 Est-ce possible .?
- Est-ce que le fait d'avoir accès à la WIFI modifie l'usage de l'ordinateur ou de la tablette, et en particulier leur usage à des fins ne relevant pas de l'apprentissage en lien avec le cours comparaison des étudiants ayant répondu ot ou osp à tablette et ordi et appartenant d'un côté au group (SV1 sv2 Sa1 M1 M2 E1 B1 et B2) par rapport au groupe (HUS1 HUS 2 HUS3 H1 H2 C1 C2C3) sur la question 7
- Est-ce que ceux qui avaient leur téléphone ou leur smartphone sur la table l'ont plus utilisé que ceux qui l'avaient dans leur sac, et pour faire quoi ? quelque soit la génération

3 Résultats selon la promotion

On crée une variable supplémentaire PROMO:

```
## p1 p2 p3
## 568 365 513
```

p = propmation, chiffre = année

Comparaison 1ère année (S1 + S2), 2ème année (S3 + S4), 3ème année (S5 + S6) comparaison de la population(HUS1, SV1, C2, B2, et M1) avec (HUS3, H1, E1, C1) avec (Sa1, SV2, HUS2 , C3, H2, B1et M2) sur Q4 et Q7.

3.1 Question Q4

La queestion 4 comporte 16 sous questions dichotomiques, correspondant aux colonnes 15 à 30. Il faut éliminer la colonne 27 qui contient du texte libre. Enfin on ajoute les colonnes 52 (age) et 53 (promotion) pour une analyse en sous-groupes. Au final le dataframe q4 comporte 15 colonnes dichotomiques + 2 colonnes servant à fabriquer des sous-groupes.

	p1	p2	p3
Q4.1	83	45	68
Q4.2	80	53	137
Q4.3	13	6	6
Q4.4	340	162	326
Q4.5	125	51	200
Q4.6	3	1	1
Q4.7	13	7	13
Q4.8	101	51	120
Q4.9	28	15	19
Q4.10	79	44	104

```

Q4.11  36   8  20
Q4.12 144  62 124
Q4.14 132 125  83
Q4.15  22  27  18
Q4.16   6   2   0

```

suite: comparaison des 3 groupes pour chaque sous-question par test chi2 après transformation des NA en '0'.

```

      0      1
1250 196

```

```

q      p1  p2  p3
0 485 320 445
1  83  45  68

```

Pearson's Chi-squared test

```

data: table(q, q4$PROMO)
X-squared = 1.05, df = 2, p-value = 0.5916

```

```

q      p1  p2  p3
0 485 320 445
1  83  45  68

```

```

q      p1      p2      p3
0 491.00968 315.52559 443.46473
1  76.99032  49.47441  69.53527

```

Alternative avec le test **prop.test**. La ligne *sample estimate* donne la proportion de 'non' dans chaque promotion.

3-sample test for equality of proportions without continuity correction

```

data: t(table(q, q4$PROMO))
X-squared = 1.05, df = 2, p-value = 0.5916
alternative hypothesis: two.sided
sample estimates:
  prop 1    prop 2    prop 3 
0.8538732 0.8767123 0.8674464

```

A la question *chercher sur internet des informations qui me manquaient au sujet du cours*, il n'y a pas de différence de comportement entre les promotions.

Pour chacune des 15 sous-questions on compare les réponses de 3 promotions par un test du chi2. Si **p-value** est supérieur à 0.05, alors il n'y a pas de différence de comportement entre les promotions pour cette sous-question. Dans la cas contraire, au moins une promotion ne se comporte pas comme les autres:

```

[1] "Q4.1"

Pearson's Chi-squared test

data:  table(q, q4$PROMO)
X-squared = 1.05, df = 2, p-value = 0.5916

[1] "Q4.2"

Pearson's Chi-squared test

data:  table(q, q4$PROMO)
X-squared = 33.8167, df = 2, p-value = 0.00000004537

[1] "Q4.3"

Pearson's Chi-squared test

data:  table(q, q4$PROMO)
X-squared = 2.0079, df = 2, p-value = 0.3664

[1] "Q4.4"

Pearson's Chi-squared test

data:  table(q, q4$PROMO)
X-squared = 34.5843, df = 2, p-value = 0.00000003091

[1] "Q4.5"

Pearson's Chi-squared test

data:  table(q, q4$PROMO)
X-squared = 77.1108, df = 2, p-value < 2.2e-16

Warning in chisq.test(table(q, q4$PROMO)): Chi-squared approximation may be
incorrect

[1] "Q4.6"

Pearson's Chi-squared test

data:  table(q, q4$PROMO)
X-squared = 0.9417, df = 2, p-value = 0.6245

[1] "Q4.7"

Pearson's Chi-squared test

data:  table(q, q4$PROMO)
X-squared = 0.3634, df = 2, p-value = 0.8338

[1] "Q4.8"

```

```

Pearson's Chi-squared test

data: table(q, q4$PROMO)
X-squared = 13.0376, df = 2, p-value = 0.001475

[1] "Q4.9"

Pearson's Chi-squared test

data: table(q, q4$PROMO)
X-squared = 1.0248, df = 2, p-value = 0.5991

[1] "Q4.10"

Pearson's Chi-squared test

data: table(q, q4$PROMO)
X-squared = 13.1483, df = 2, p-value = 0.001396

[1] "Q4.11"

Pearson's Chi-squared test

data: table(q, q4$PROMO)
X-squared = 9.5534, df = 2, p-value = 0.008424

[1] "Q4.12"

Pearson's Chi-squared test

data: table(q, q4$PROMO)
X-squared = 9.6521, df = 2, p-value = 0.008018

[1] "Q4.14"

Pearson's Chi-squared test

data: table(q, q4$PROMO)
X-squared = 38.7471, df = 2, p-value = 0.000000003856

[1] "Q4.15"

Pearson's Chi-squared test

data: table(q, q4$PROMO)
X-squared = 8.521, df = 2, p-value = 0.01411

Warning in chisq.test(table(q, q4$PROMO)): Chi-squared approximation may be
incorrect

[1] "Q4.16"

Pearson's Chi-squared test

```



```
data: table(q, q4$PROMO)
X-squared = 5.4671, df = 2, p-value = 0.06499
```

Même analyse en utilisant le **test exact de Fisher** qui donne un résultat plus précis que la chi2 en cas d'effectifs faibles:

```
[1] "Q4.1"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.6038
alternative hypothesis: two.sided
```

```
[1] "Q4.2"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.00000008304
alternative hypothesis: two.sided
```

```
[1] "Q4.3"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.3576
alternative hypothesis: two.sided
```

```
[1] "Q4.4"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.00000003458
alternative hypothesis: two.sided
```

```
[1] "Q4.5"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value < 2.2e-16
alternative hypothesis: two.sided
```

```
[1] "Q4.6"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.8525
alternative hypothesis: two.sided
```

[1] "Q4.7"

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.8561
alternative hypothesis: two.sided
```

[1] "Q4.8"

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.001524
alternative hypothesis: two.sided
```

[1] "Q4.9"

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.6171
alternative hypothesis: two.sided
```

[1] "Q4.10"

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.00164
alternative hypothesis: two.sided
```

[1] "Q4.11"

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.007551
alternative hypothesis: two.sided
```

[1] "Q4.12"

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.006745
alternative hypothesis: two.sided
```

[1] "Q4.14"

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
p-value = 0.000000005307
```

```
alternative hypothesis: two.sided
```

```
[1] "Q4.15"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
```

```
p-value = 0.01968
```

```
alternative hypothesis: two.sided
```

```
[1] "Q4.16"
```

Fisher's Exact Test for Count Data

```
data: table(q, q4$PROMO)
```

```
p-value = 0.04104
```

```
alternative hypothesis: two.sided
```

Ne pas tenir compte de cette remarque

Modèle plus complexe (utilité ?). [Référence](#) fiche 61 et 62.

```
# f <- glm(q1 ~ q4$PROMO, family = "binomial")
# f
# summary(f)
```

3.2 Question Q7

La question 7 comporte 15 sous questions de Q7.1 à Q7.16 (colonnes 33 à 48). On retire la colonne 13 qui est du texte libre. On ajoute la colonne PROMO (promotion) pour l'analyse en sous-groupe.

	p1	p2	p3
Q7.1	22	18	23
Q7.2	19	10	26
Q7.3	13	6	20
Q7.4	25	7	38
Q7.5	0	2	5
Q7.6	3	1	1
Q7.7	17	6	19
Q7.8	2	2	1
Q7.9	22	6	14
Q7.10	24	30	37
Q7.11	191	89	88
Q7.12	28	18	7
Q7.14	86	72	76
Q7.15	252	129	259
Q7.16	9	4	1

REMARQUE: les lignes pour lesquelles un des effectif est inférieur à 5 donnent des résultats douteux au test du chi2.

Pour chacune des 15 sous-questions on compare les réponses de 3 promotions par un test du chi2. Si **p-value** est supérieur à 0.05, alors il n'y a pas de différence de comportement entre les promotions pour cette sous-question. Dans la cas contraire, au moins une promotion ne se comporte pas comme les autres:

```
[1] "Q7.1"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 0.6278, df = 2, p-value = 0.7306
```

```
[1] "Q7.2"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 3.6977, df = 2, p-value = 0.1574
```

```
[1] "Q7.3"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 4.7259, df = 2, p-value = 0.09414
```

```
[1] "Q7.4"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 14.3437, df = 2, p-value = 0.0007679
```

```
Warning in chisq.test(table(q, q7$PROMO)): Chi-squared approximation may be
incorrect
```

```
[1] "Q7.5"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 5.3566, df = 2, p-value = 0.06868
```

```
Warning in chisq.test(table(q, q7$PROMO)): Chi-squared approximation may be
incorrect
```

```
[1] "Q7.6"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 0.9417, df = 2, p-value = 0.6245
```

```
[1] "Q7.7"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 3.2345, df = 2, p-value = 0.1984
```

```
Warning in chisq.test(table(q, q7$PROMO)): Chi-squared approximation may be incorrect
```

```
[1] "Q7.8"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 0.7723, df = 2, p-value = 0.6797
```

```
[1] "Q7.9"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 4.003, df = 2, p-value = 0.1351
```

```
[1] "Q7.10"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 7.1495, df = 2, p-value = 0.02802
```

```
[1] "Q7.11"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 38.8441, df = 2, p-value = 0.000000003674
```

```
[1] "Q7.12"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 11.9195, df = 2, p-value = 0.002581
```

```
[1] "Q7.14"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 4.5408, df = 2, p-value = 0.1033
```

```
[1] "Q7.15"
```

```
Pearson's Chi-squared test
```

```
data: table(q, q7$PROMO)
X-squared = 19.8318, df = 2, p-value = 0.00004938
```

```
Warning in chisq.test(table(q, q7$PROMO)): Chi-squared approximation may be incorrect
```

```
[1] "Q7.16"
```

Pearson's Chi-squared test

```
data: table(q, q7$PROMO)
X-squared = 5.5114, df = 2, p-value = 0.06356
```

Même analyse en utilisant le **test exact de Fisher** qui donne un résultat plus précis que la chi2 en cas d'effectifs faibles:

```
[1] "Q7.1"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.7252
alternative hypothesis: two.sided
```

```
[1] "Q7.2"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.1851
alternative hypothesis: two.sided
```

```
[1] "Q7.3"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.11
alternative hypothesis: two.sided
```

```
[1] "Q7.4"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.000556
alternative hypothesis: two.sided
```

```
[1] "Q7.5"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.04439
alternative hypothesis: two.sided
```

```
[1] "Q7.6"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.8525
alternative hypothesis: two.sided
```

```
[1] "Q7.7"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.193
alternative hypothesis: two.sided
```

```
[1] "Q7.8"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.7437
alternative hypothesis: two.sided
```

```
[1] "Q7.9"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.1436
alternative hypothesis: two.sided
```

```
[1] "Q7.10"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.02268
alternative hypothesis: two.sided
```

```
[1] "Q7.11"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.000000003063
alternative hypothesis: two.sided
```

```
[1] "Q7.12"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.001148
alternative hypothesis: two.sided
```

```
[1] "Q7.14"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.1091
alternative hypothesis: two.sided
```

```
[1] "Q7.15"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.00004577
alternative hypothesis: two.sided
```

```
[1] "Q7.16"
```

Fisher's Exact Test for Count Data

```
data: table(q, q7$PROMO)
p-value = 0.04678
alternative hypothesis: two.sided
```

4 Information de session

Informations pour le chapitre matériel et méthode.

R version 3.1.3 (2015-03-09)
Platform: x86_64-apple-darwin13.4.0 (64-bit)
Running under: OS X 10.10.3 (Yosemite)

locale:

```
[1] fr_FR.UTF-8/fr_FR.UTF-8/fr_FR.UTF-8/C/fr_FR.UTF-8/fr_FR.UTF-8
```

attached base packages:

```
[1] stats      graphics  grDevices  utils      datasets  methods   base
```

other attached packages:

```
[1] knitr_1.10.5    xtable_1.7-4    stringr_1.0.0    epicalc_2.15.1.0
[5] nnet_7.3-9      MASS_7.3-40     survival_2.38-1  foreign_0.8-63
```

loaded via a namespace (and not attached):

```
[1] digest_0.6.8    evaluate_0.7     formatR_1.2
[4] htmltools_0.2.6 magrittr_1.5     rmarkdown_0.5.3.2
[7] splines_3.1.3   stringi_0.4-1    tools_3.1.3
[10] yaml_2.1.13
```

To cite R in publications use:

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

A BibTeX entry for LaTeX users is

```
@Manual{,
  title = {R: A Language and Environment for Statistical Computing},
  author = {{R Core Team}},
  organization = {R Foundation for Statistical Computing},
  address = {Vienna, Austria},
  year = {2015},
  url = {http://www.R-project.org/},
}
```

We have invested a lot of time and effort in creating R, please
cite it when using it for data analysis. See also
'citation("pkgname")' for citing R packages.