

**Analyse factorielle et K-means clustering**  
**pour campagne de collecte de fonds**

## **Analyse factorielle sur profil de donateurs en campagne de dons**

En premier lieu, nous avons entamé une analyse pour valider la qualité des données contenues dans *TestPersonnalite.csv*. Celle-ci nous a permis de constater que les questions évaluant la personnalité étaient tirées du Big-Five<sup>1</sup>. Cela nous a énormément simplifié la tâche car nous savions désormais le nombre de facteurs à évaluer ainsi que le type de rotation à employer. Effectivement, le Big-Five comporte 5 composantes, ce qui devrait logiquement constituer la solution optimale pour les données dont nous disposons. Concernant la rotation, nous savons que les composantes du Big-Five sont plus ou moins corrélées entre elles. Il sera donc plus utile d'utiliser une rotation oblique qu'une rotation orthogonale comme varimax. En l'occurrence, nous avons utilisé promax, ce qui a donné de très bons résultats. En effet, nous avons extrait cinq facteurs dont aucune n'excédait un *cross-loading* de 0,45. En d'autres mots, le chargement des items sur leurs facteurs respectifs était très élevé et très faible sur les facteurs qui n'étaient pas les leurs.

Il est toutefois étonnant de constater que les items ne chargent pas les facteurs qu'ils devraient charger selon la littérature scientifique. Par exemple, les adjectifs du Facteur2 appartiennent tous à des composantes différentes du Big-Five.

Néanmoins, la validité interne des variables extraites de l'analyse factorielle est excellente (Nunnally, 1967)<sup>2</sup>, puisque les scores des alphas de Cronbach se distribuent entre [0.91; 0.97].

Par précaution, nous avons ajusté le modèle pour connaître le nombre de facteurs optimal. Nous avons évalué les indices pour chacun des 6 facteurs et la solution à 5 facteurs est ressortie gagnante avec le plus petit BIC, mais aussi comme modèle le

k	AIC	BIC	pval	npar	heywood
1	5511.71878	5652.2571	< 2.2e-16	30	oui
2	-41.10728	165.0156	< 2.2e-16	44	oui
3	-4846.11905	-4579.0962	< 2.2e-16	57	oui
4	-10066.66084	-9743.4226	< 2.2e-16	69	oui
5	-16265.54430	-15890.7754	5.576e-05	80	non
6	-16285.87125	-15864.2562	0.04104	90	non

---

<sup>1</sup> <https://fetzer.org/sites/default/files/images/stories/pdf/selfmeasures/PersonalityBigFiveInventory.pdf>

<sup>2</sup> Nunnally, J. C. (1967). Psychometric theory. McGraw-Hill.

plus parcimonieux n'étant pas un cas de Quasi-Heywood. Par ailleurs, nous avons choisi d'ignorer le fait que la solution à 6 facteurs était celle avec le plus petit AIC car cet indice est généralement moins sévère que le BIC.

Compte tenu de la convergence de ces indices qui semblent démontrer que ces facteurs représentent adéquatement 5 variables latentes au sein des données, nous pourrions créer des variables. Voici les variables créées et les items de mesure les composant:

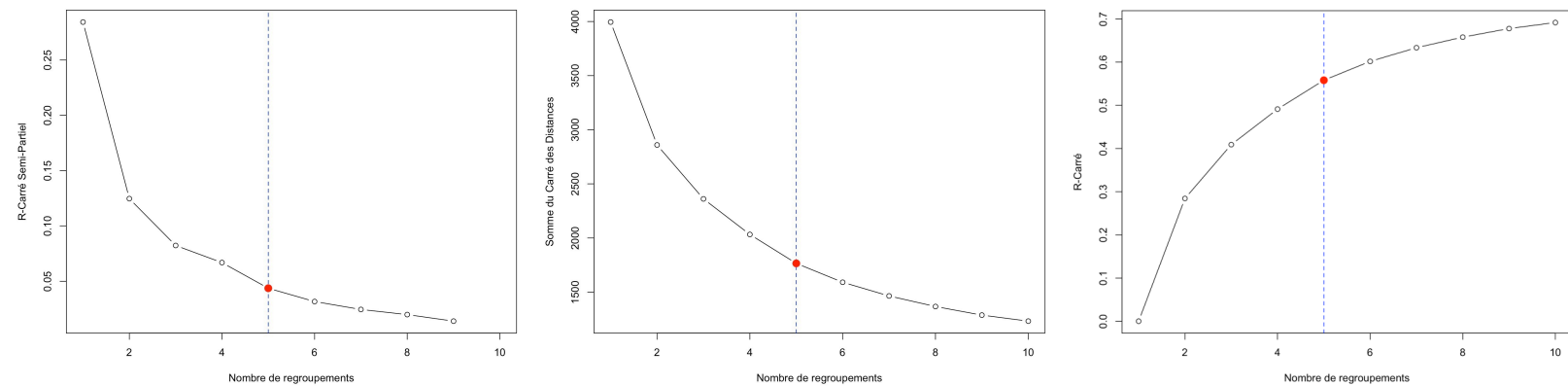
- **Extraversion** = Talkative, Outgoing, Forgiving
- **Conscientiousness** = Thorough, Efficient, Energetic
- **Agreeableness** = Helpful\_unselfish, Considerate, Cooperative
- **Openness** = Relaxed (Full\_of\_ideas), Emotionally\_stable  
(Quick\_to\_understand\_things), Sophisticated\_in\_arts
- **Neuroticism** = Easily\_distracted, Moody (Easily\_irritated), Curious

## Analyse de regroupement

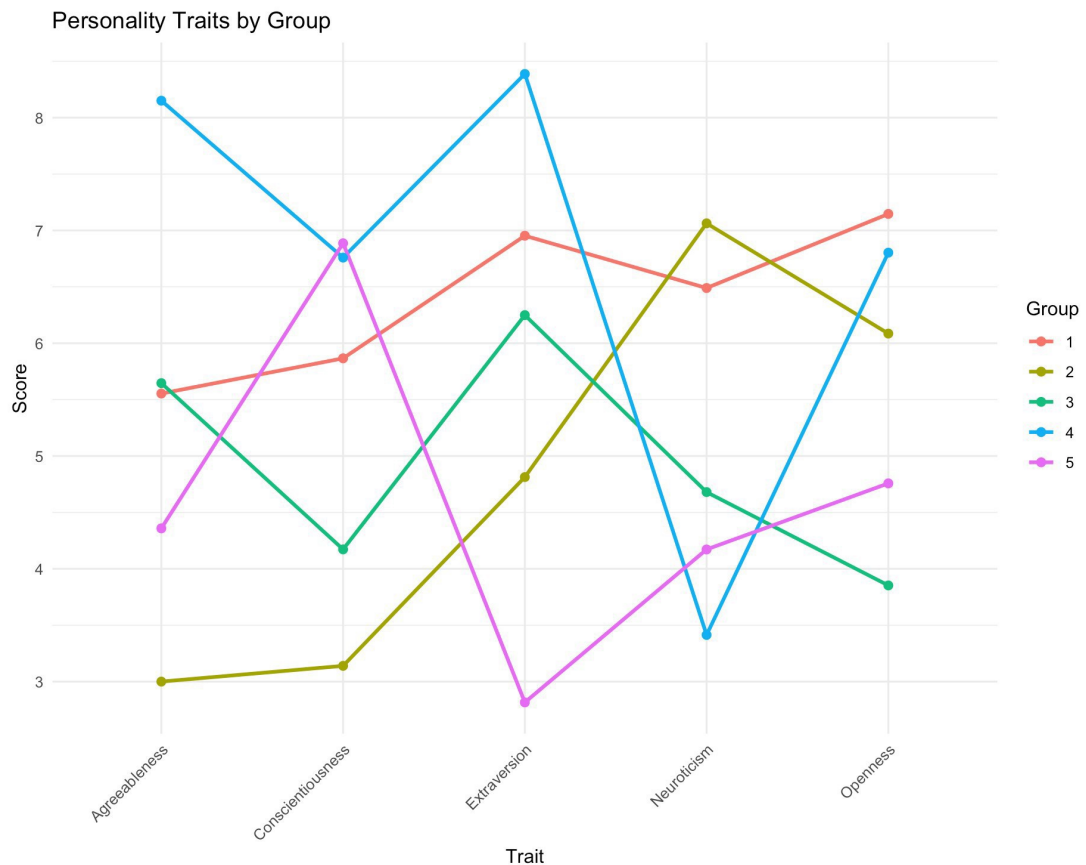
Nous avons ensuite mené une analyse de regroupements avec l'algorithme k-means pour segmenter l'échantillon à partir des échelles décrites à la question précédente. Tout d'abord, nous avons défini un seed pour assurer la reproductibilité de nos résultats. Étant donné que toutes les variables utilisées pour l'analyse étaient sur la même

échelle, nous ne les avons pas standardisé. Afin de trouver le nombre de regroupements correspondant le mieux à nos données, nous avons exécuté une analyse de K-means pour un nombre de regroupements de  $k=1$  à 10 et avons comparé le  $R_2$ , le  $R_2$  semipartiel, le BIC (la somme totale du carré des distances intragroupes pour chaque nombre de regroupements).

Tous les indices semblaient suggérer que le nombre de regroupements optimal dépassait la limite supérieure de 10, préalablement définie. Cependant, nous avons constaté un point de rendement décroissant, où l'ajout d'un segment n'améliorait pas fortement la qualité du modèle. En comparant les différents tracés des graphiques, nous avons conclu que la solution à 5 groupes était la plus parcimonieuse, tout en étant celle qui possédait les meilleures indices d'ajustement.



## Interprétation:



### Groupe 1 : Les Idéalistes Extravertis

- Extraversion : 6,95 (Élevée)
- Conscienciosité : 5,87 (Relativement élevée)
- Amabilité : 5,55 (Modérée)
- Ouverture : 7,15 (Élevée)
- Névrosisme : 6,49 (Relativement élevé)

Comportement : Les individus du groupe 1 sont extravertis et ouverts aux nouvelles expériences, avec un niveau modéré de conscienciosité et d'amabilité. Cependant, leur névrosisme relativement élevé indique qu'ils peuvent connaître des hauts et des bas émotionnels. Ils sont probablement créatifs et enthousiastes, mais peuvent avoir des difficultés avec la stabilité émotionnelle. Ces traits les rendent donc idéalistes mais potentiellement réactifs.

### Groupe 2 : Les Réservés Inquiets

- Extraversion : 4,81 (Faible)
- Conscienciosité : 3,14 (Faible)
- Amabilité : 3,00 (Faible)
- Ouverture : 6,09 (Relativement élevée)
- Névrosisme : 7,06 (Élevé)

Comportement : Le groupe 2 est plutôt introverti et émotionnellement réactif, avec un faible niveau de conscienciosité et d'amabilité. Bien qu'ils soient ouverts à de nouvelles idées, leur névrosisme élevé et leur faible amabilité peuvent entraîner des défis interpersonnels ou psychosociaux. Ils peuvent être analytiques, mais ont du mal à maintenir des engagements sociaux ou liés au travail à cause de leur faibles compétences interpersonnelles.

### **Groupe 3 : Les Indépendants Pragmatistes**

- Extraversion : 6,25 (Modérée)
- Conscienciosité : 4,17 (Relativement faible)
- Amabilité : 5,65 (Relativement élevée)
- Ouverture : 3,85 (Faible)
- Névrosisme : 4,68 (Relativement faible)

Comportement : Le groupe 3 montre une sociabilité modérée et une stabilité émotionnelle moyenne, avec une approche pratique et indépendante en raison de leur faible ouverture. Leur conscienciosité et amabilité modérées indiquent qu'ils sont généralement coopératifs et fiables, mais pourraient ne pas exceller dans des situations nécessitant une grande créativité ou adaptabilité.

### **Groupe 4 : Les Leaders Dynamiques**

- Extraversion : 8,39 (Très élevée)
- Conscienciosité : 6,76 (Élevée)
- Amabilité : 8,15 (Très élevée)
- Ouverture : 6,80 (Élevée)
- Névrosisme : 3,41 (Faible)

Comportement : Les résultats du groupe 4 suggèrent un fort potentiel de leadership. Ils sont très sociables, émotionnellement stables et extrêmement agréables, ce qui en fait des leaders charismatiques et empathiques. Leur conscienciosité élevée et leur ouverture montrent qu'ils sont organisés, motivés et ouverts à l'innovation. Ce groupe excelle probablement dans des rôles d'équipe et de leadership.

### **Groupe 5 : Les Réalistes Analytiques**

- Extraversion : 2,82 (Faible)
- Conscienciosité : 6,89 (Élevée)
- Amabilité : 4,36 (Relativement faible)
- Ouverture : 4,76 (Relativement faible)
- Névrosisme : 4,17 (Relativement faible)

Comportement : Le groupe 5 a tendance à être réservé et orienté vers les tâches, avec une conscienciosité élevée et une amabilité relativement faible. Leur faible ouverture et extraversion indiquent qu'ils préfèrent la routine et la structure à la nouveauté et aux interactions sociales. Cependant, leur stabilité émotionnelle les rend fiables et constants, excellant dans des rôles nécessitant précision et réflexion analytique.

## Généralisation sur la population globale

Nous avons créé une base de donnée en faisant correspondre les ID des individus de la base de données *TestPersonnalite.csv* ainsi que ceux des bases de données *DonationHistory.csv*, *MembersList.csv*. À cette base de données, nous avons rajouté les informations de la base de données *NewsLetterRead.csv* en faisant correspondre le email des participants. Nous n'avons pas utilisé les variables contenues dans un autre fichier *SocialNetworkUsage.csv*, car nous avons jugé que le lien unissant ces données aux données des autres bases de données était trop ténu pour être exploitable. Effectivement, même s'il est possible de faire correspondre les courriels des individus d'une base de donnée à l'autre, leurs noms ne correspondent pas. Essayer de décrire des individus différents pourrait donc biaiser nos analyses.

Pour simplifier l'interprétation des variables contenues dans les bases de données, nous en avons créé de nouvelles. Effectivement, nous avons fait la somme des dons de chaque individus pour n'avoir qu'une seule variable: `dons_total`. De même, nous avons fait la somme du nombre de fois où des individus ont lu l'infolettre, ce qui nous a permis de créer la variable: `total_reads`.

Au total, nous avons utilisé 8 variables :

1. **Woman**: Le fait d'être une femme ou non.
2. **Age**: L'âge des individus.
3. **Salary**: Le salaire annuel des individus.
4. **Education**: Le plus haut niveau d'éducation atteint par les individus.
5. **City**: Le type de milieu urbain dans lequel habitent les membres de la base de données.
6. **Joined**: L'année d'adhésion des membres à la fondation.
7. **total\_reads**: Le nombre total de fois que les membres de la base de données ont ouvert l'infolettre pendant l'année 2023.
8. **dons\_total**: Le montant total des donations des membres de la base de données, de 2013 à 2022.

Dans la section qui suit, nous décrirons chaque segment au regard de son score à chacune des variables ci-haut. Ensuite, un court paragraphe résumera comment chaque segment se distingue d'un autre. Enfin, dans un troisième paragraphe, nous tenterons d'émettre des hypothèses sur la manière dont les traits de personnalité ont influencé les scores des individus sur les variables qui distinguent leur segment d'appartenance.



<b>Variable</b>	<b>Groupe 1</b> (n=210)	<b>Groupe 2</b> (n=162)	<b>Groupe 3</b> (n=149)	<b>Groupe 4</b> (n=149)	<b>Groupe 5</b> (n=130)
<b>Woman</b>					
<i>0</i>	45% (n=94)	45% (n=73)	56% (n=83)	47% (n=70)	43% (n=56)
<i>1</i>	55% (n=116)	55% (n=89)	44% (n=66)	53% (n=79)	57% (n=74)
<b>Age</b>	M = 43 ÉT = 18	M = 48 ÉT = 19	M = 48 ÉT = 19	M = 48 ÉT = 19	M = 44 ÉT = 18
<b>Salary</b>	M = 71,843 ÉT = 66,283	M = 65,769 ÉT = 58,547	M = 71,291 ÉT = 66,046	M = 65,161 ÉT = 61,313	M = 69,137 ÉT = 69,546
<b>Education</b>					
<i>Elementary</i>	0% (n=0)	0.6% (n=1)	2.0% (n=3)	4.7% (n=7)	3.1% (n=4)
<i>High School</i>	26% (n=54)	22% (n=36)	32% (n=48)	39% (n=58)	42% (n=54)
<i>University / College</i>	74% (n=156)	77% (n=125)	66% (n=98)	56% (n=84)	55% (n=72)
<b>City</b>					
<i>City</i>	20% (n=43)	34% (n=55)	21% (n=31)	36% (n=53)	24% (n=31)
<i>Downtown</i>	19% (n=40)	17% (n=28)	21% (n=31)	23% (n=35)	3.1% (n=4)
<i>Rural</i>	26% (n=55)	13% (n=21)	19% (n=29)	9.4% (n=14)	55% (n=72)
<i>Suburban</i>	34% (n=72)	36% (n=58)	39% (n=58)	32% (n=47)	18% (n=23)
<b>Joined</b>	M = 2,018 ÉT = 3	M = 2,017 ÉT = 3	M = 2,017 ÉT = 3	M = 2,018 ÉT = 3	M = 2,018 ÉT = 3
<b>total_reads</b>	M = 1.20 ÉT = 1.20	M = 4.18 ÉT = 2.23	M = 0.38 ÉT = 0.66	M = 1.93 ÉT = 1.52	M = 1.55 ÉT = 1.48
<b>dons_total</b>	M = 18 ÉT = 38	M = 78 ÉT = 118	M = 5 ÉT = 20	M = 34 ÉT = 66	M = 21 ÉT = 55

### Groupe 1:

La répartition homme-femme est équilibrée (45% hommes, 55% femmes). Bien que ça soit le groupe le plus jeune (M = 43, ÉT = 18) c'est celui dont les membres ont le deuxième meilleur salaire moyen (M = 71,843, ÉT = 66,283), ce qui pourrait s'expliquer par leur niveau éducation élevée. Effectivement, le groupe 1 est composé d'une très forte proportion de diplômés universitaires (74%) et tous ses membres possède au moins un diplôme d'études secondaires. Ils habitent dans des milieux urbains variés, avec des proportions équilibrées entre la ville (20%), la banlieue (34%), le centre-ville (19%), et les

milieux ruraux (26%). En outre, ils semblent peu intéressés par les activités de la fondation (total\_reads M = 1.20) et y ont fait relativement peu de dons (M = 18; ÉT = 38).

En conclusion, le Groupe 1 se distingue par un profil équilibré, jeune et très diplômé, mais avec une participation relativement faible (lectures et dons limités).

Le segment 1 est celui que nous avons nommé "Les Idéalistes Extravertis". Il se distingue par des scores relativement élevés d'ouverture et d'extraversion de ses membres. On pourrait supposer que ces deux attributs leur ont insufflé l'envie de poursuivre des études universitaires.

### **Groupe 2:**

Comme le groupe 1, la répartition homme-femme est équilibrée (45% hommes, 55% femmes), mais ils sont plus âgés (M = 48; ÉT = 19). Malgré leur salaire qui figure parmi les plus bas, c'est le groupe le plus généreux en terme de dons (M = 78; ÉT = 118), peut-être parce qu'ils sont aussi les plus intéressés par les activités de la fondation, ce que l'on peut constater par leur intérêt envers la lecture de l'infolettre (M = 4,18; ÉT = 2,23). Ce sont effectivement de grands lecteurs, ce qui n'est pas étonnant venant du segment avec la plus grande proportion de diplômés universitaires (77%). Plus souvent qu'autrement, ils habitent en ville (34%) ou en banlieue (36%). Il semble que la proximité de la ville soit importante pour eux puisqu'on les retrouve peu en milieu rural (13%).

En conclusion, le Groupe 2 est caractérisé par une population plus âgée, très diplômée et vivant principalement en ville/banlieue. Ce groupe est le plus engagé (lectures et dons élevés).

Le segment 2 est celui que nous avons nommé "Les Réservés Inquiets". Il est étonnant de constater que ce sont les plus grands donateurs, étant donné qu'ils ont les scores d'amabilité les plus bas et les scores de névrosisme les plus élevés. Nous avons émis l'hypothèse que l'interaction de ces traits était susceptible de mener à des problématiques interpersonnelles. Peut-être que les dons sont une manière pour eux de garder une bonne conscience par rapport au manque d'amabilité dont ils font preuve.

### **Groupe 3:**

Le troisième groupe rassemble la plus grande proportion d'hommes (56%), mais se trouve aussi à être le deuxième groupe à le salaire le plus élevé (M = 71,291). Malgré leurs revenus élevés, c'est le groupe le moins généreux en terme de donations (M = 5; ÉT = 20), ce qui n'est pas étonnant au regard du nombre de fois peu élevé où ils ont ouvert l'infolettre cette année (M = 0,38; ÉT = 0,66). Cela contraste avec le niveau d'éducation assez élevé de ce groupe, où seulement 2% des individus ne possèdent pas de diplôme d'études secondaires, par rapport à 32% en possédant un et où 66% des

individus ont un diplôme universitaire. Par ailleurs, c'est le groupe dont le lieu d'habitation est réparti de la manière la plus homogène à travers les différents milieux urbains.

En conclusion, le Groupe 3 est composé principalement d'hommes, plus âgés, avec un faible niveau d'engagement malgré un salaire et une éducation élevés.

Le segment 3 est celui que nous avons nommé "Les Indépendants Pragmatistes". Leur indépendance les rend probablement moins susceptible de considérer faire un don à une fondation caritative. En amont, leur faible ouverture aux nouvelles expériences exacerbe probablement cette réticence, les rendant également insensibles aux campagnes marketing les incitant à faire un don, ce qui explique pourquoi ils sont le groupe le moins engagé auprès de la fondation.

#### **Groupe 4:**

Le sexe des individus de ce groupe est réparti de manière équilibrée, mais ils figurent parmi les groupes dont l'âge moyen est le plus élevé. C'est aussi le groupe avec la plus grande proportion d'individus dont le niveau d'études le plus élevé est le primaire (4,7%), ce qui explique peut-être pourquoi c'est aussi le groupe avec le salaire moyen le plus bas ( $M = 65,161$ ). Ce groupe semble avoir les milieux ruraux en dédain, ses membres y habitant peu (9,4%) comparativement aux autres milieux urbains dans lesquels ils habitent en proportion assez égale. Par ailleurs, ce groupe semble plutôt engagé dans la fondation, lisant son infolettre en moyenne environ deux fois par année et ayant donné en moyenne 34\$ depuis leur adhésion à la fondation.

En conclusion, le Groupe 4 présente un profil moins diplômé, avec des revenus plus faibles. Il est surtout urbain et modérément engagé.

Le segment 4 est celui que nous avons nommé "Les Leaders Dynamiques". Ses membres sont très aimables et très extravertis. Ils cherchent probablement le contact humain, ce qui les pousse à préférer les milieux urbains aux milieux ruraux. Compte tenu du nom que nous avons attribué au segment, nous nous serions attendu à des individus avec des emplois de gestion et des salaires élevés. Cependant, compte tenu de l'âge avancé de ce segment, peut-être qu'une partie de ces gens est à la retraite, ce qui abaisserait la moyenne des revenus du groupe.

#### **Groupe 5:**

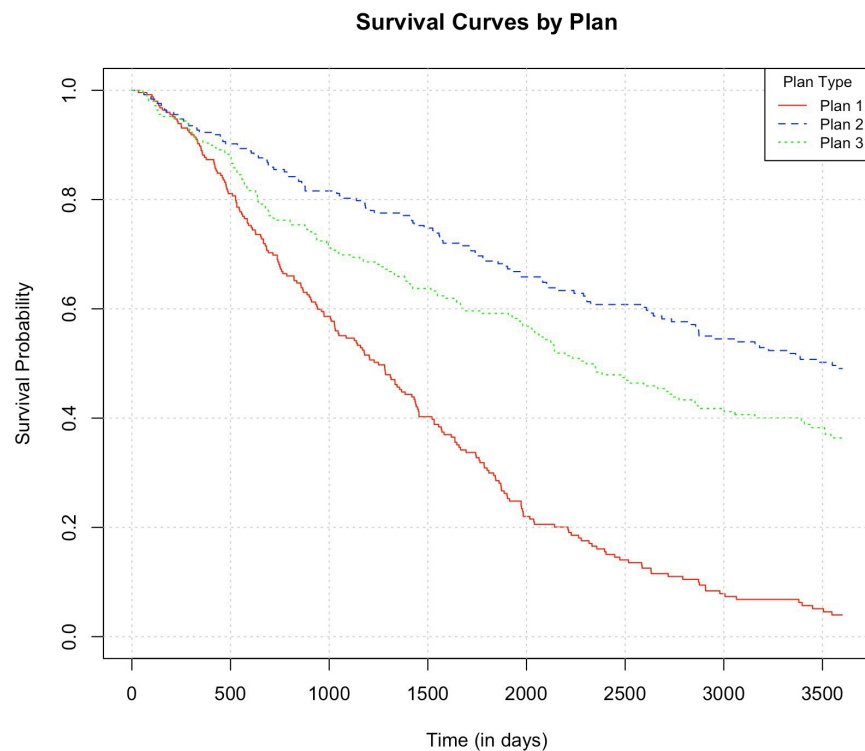
Le groupe 5 est celui dans lequel il y a le plus de femmes (57%). Comme le premier groupe, il se distingue par le jeune âge de ses membres ( $M = 44$ ) et leur salaire relativement élevé ( $M = 69,137$ ), mais très variable ( $ÉT = 69,546$ ). C'est toutefois le segment qui a le moins haut taux de diplômés universitaire (55%), tout en ayant le taux le plus élevé de diplômés du secondaire. Le milieu rural constitue le lieu de vie de 55% des membres de ce groupe, une proportion bien supérieure à celle des autres groupes. Ils semblent effectivement préférer un milieu de vie calme, puisque seulement 3,1% des individus de ce groupe habitent au centre-ville, la proportion la plus faible de tous les

groupes. De plus, ils sont moyennement engagé avec la fondation, ne lisant son infolettre qu'en moyenne 1,55 fois par année et ayant donné en moyenne 21\$ depuis leur adhésion à la fondation.

En conclusion, le Groupe 5 est distinct par sa surreprésentation féminine, sa forte ruralité, et une éducation moins élevée. Il est faiblement engagé malgré un salaire relativement élevé.

Le segment 5 est celui que nous avons nommé "Les Réalistes Analytiques". Il se distingue par sa grande introversion. Il n'est donc pas étonnant de les voir préférer les milieux ruraux aux milieux urbains. Il est toutefois étonnant de constater leur faible niveau d'étude, une forte conscienciosité en étant généralement un facteur prédicteur. Peut-être que leur faible ouverture aux nouvelles expériences et leur préférence pour la routine les a poussé vers des emplois plus traditionnels, requérant moins d'études.

**Question 2 a:** i) Voir le graphique ci-bas.



ii) Les probabilités qu'une pompe soit encore fonctionnelle après 1800 jours sont consignées dans le très beau tableau ci-bas.

Survival at 1800 Days by Plan						
Includes survival probabilities and confidence intervals						
Plan	Time (Days)	Number at Risk	Number of Events	Survival Probability	Lower 95% CI	Upper 95% CI
Plan=1	1800	66	161	30.4%	25.0%	37.1%
Plan=2	1800	145	72	68.7%	63.0%	75.0%
Plan=3	1800	127	97	59.2%	53.2%	65.8%

Pour le plan 1, la probabilité qu'une pompe soit encore en activité après 1800 jours est de 30,4%.

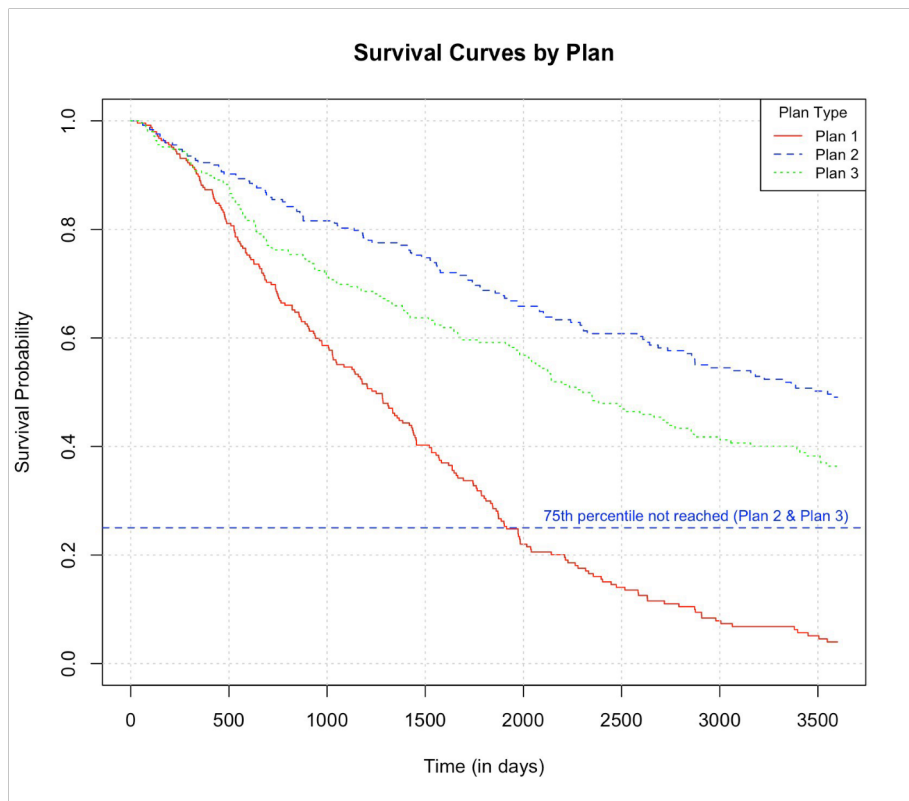
Pour le plan 2, la probabilité qu'une pompe soit encore en activité après 1800 jours est de 68,7%.

Pour le plan 3, la probabilité qu'une pompe soit encore en activité après 1800 jours est de 59,2%.

iii) Ci-bas sont représentés les quartiles du temps de vie pour les trois plans de maintenance.

Plan	quantile.25	quantile.50	quantile.75
Plan 1	605	1250	1914
Plan 2	1481	3550	NA
Plan 3	877	2303	NA

iv) Nous sommes effectivement incapable d'obtenir le dernier quartile pour les plans 2 et 3. Effectivement, ce quartile représente le temps où 75% des pompes ont faillies. Or, puisqu'il reste plus de 25% du nombre de pompes de ces deux plans, elle n'ont pas encore atteint le moment où 75% des pompes ont fait défaut. Nous avons produit un graphique qui décrit bien ce phénomène.



Nous avons commencé par ajuster le modèle de Cox. Puisque le test du rapport de vraisemblance a démontré que l'effet de la variable plan sur la durée de survie était statistiquement significatif, au seuil de  $p < 0,05$ .

ID	loglik	Chisq	Df	Pr(> Chi )
NULL	-2783.926	NA	NA	NA
Plan	-2721.744	124.3636	2	0.00000000000000000000000009880502

i) Les pompes qui bénéficient d'un plan de maintenance aux 6 mois (Plan 2) voient le risque de subir une défaillance à tout moment être diminuée d'environ 72% ( $1 - 0,282$ ) par rapport aux pompes qui ne bénéficient d'aucune maintenance préventive (Plan 1). Cet effet est par ailleurs statistiquement significatif étant donné que l'intervalle de confiance  $[0,22; 0,36]$  ne traverse pas 1 et que la valeur  $p$  est inférieure à 0,05.

Characteristic	HR <sup>†</sup>	95% CI <sup>†</sup>	p-value
Plan			
1	—	—	
2	0.28	0.22, 0.36	<0.001
3	0.42	0.34, 0.52	<0.001

<sup>†</sup> HR = Hazard Ratio, CI = Confidence Interval

ii) Les pompes qui bénéficient d'un plan de maintenance aux 12 mois (Plan 3) voient le risque de subir une défaillance à tout moment être diminuée d'environ 58% ( $1 - 0,42$ ) par rapport aux pompes qui ne bénéficient d'aucune maintenance préventive (Plan 1). Cet effet est par ailleurs statistiquement significatif étant donné que l'intervalle de confiance  $[0,34; 0,52]$  ne traverse pas 1 et que la valeur  $p$  est inférieure à 0,05.

iii) Cette question demande de calculer l'effet sur le risque instantané au plan de maintenance aux 6 mois par rapport aux plans de maintenance aux 6 mois. Nous avons pris pour acquis qu'il s'agissait d'une erreur et que nous devions plutôt calculer l'effet sur le risque instantané d'un plan de maintenance aux 12 mois par rapport au plan de maintenance systématique aux 6 mois.

Ainsi, les pompes qui bénéficient d'un plan de maintenance systématique aux 6 mois (Plan 2) voient le risque de subir une défaillance à tout moment être diminué d'environ 32% ( $1 - 0,68$ ) par rapport aux pompes qui bénéficient d'une maintenance préventive aux 12 mois (Plan 3). Cet effet est par ailleurs statistiquement significatif étant donné que l'intervalle de confiance  $[0,53; 0,87]$  ne traverse pas 1 et que la valeur  $p$  est inférieure à 0,05.

Characteristic	HR <sup>†</sup>	95% CI <sup>†</sup>	p-value
Plan			
3	—	—	
1	2.40	1.93, 2.98	<0.001
2	0.68	0.53, 0.87	0.002

<sup>†</sup> HR = Hazard Ratio, CI = Confidence Interval

En conclusion, il semble que plus les activités de maintenance sont faites de manières rapprochées, moins les pompes sont à risque de subir un bris. CEPENDANT, ces résultats devraient être interprétés avec prudence. Effectivement, le postulat de risques proportionnels n'a pas été respecté. On peut effectivement observer une tendance où la fonction de risque semble augmenter en fonction du temps dans le graphique ci-bas. En outre, le test d'hypothèse ci-bas nous permet de rejeter l'hypothèse nulle, suggérant que les risques sont stables dans le temps. Nous pouvons donc accepter l'hypothèse alternative suggérant que les risques sont non-proportionnels. Il faudrait donc rajouter un terme d'interaction à la variable Plan ou stratifier l'analyse en fonction de celle-ci.

Cox_table	chisq	df	p
Plan	19.91912	2	0.00004727354
Global	19.91912	2	0.00004727354

