



UNIVERSITÉ
CAEN
NORMANDIE



Analyse de données d'expérimentations sur l'élection présidentielle française de 2022



Stage effectué du 18 avril 2022 au 17 juin 2022

Tutrice de stage : Mme Isabelle LEBON

Superviseur académique : M. Christophe CHESNEAU

Diplôme préparé : Master 1 Statistiques appliquées et analyse décisionnelle

Laboratoire d'accueil : Centre de recherche en économie et gestion, 1 ESP de la Paix,
14000 Caen

Remerciements

Je tiens tout d'abord à remercier l'université de Caen Normandie. L'écoute et l'accompagnement dont j'ai bénéficié m'ont permis de trouver rapidement un stage, dans le but d'affiner mon projet professionnel et de valider mon Master 1 en statistiques appliquées et analyse décisionnelle.

Je pense également à Monsieur Fabien Moizeau, directeur du CREM, qui m'a accueilli au sein de son laboratoire.

À ce titre, je souhaiterais remercier tout particulièrement Madame Isabelle Lebon, professeur des universités et tutrice de stage, qui m'a épaulé et conseillé et qui m'a surtout transmis son expertise en économie.

Je souhaite également remercier Monsieur Frédéric Chantreuil, professeur à l'université de Nouvelle-Calédonie pour m'avoir apporté les connaissances théoriques sur le calcul des inégalités en Nouvelle-Calédonie, ainsi que Kevin Fourrey, Maître de conférences à l'UPEC, pour avoir pris le temps de m'expliquer le fonctionnement de l'algorithme de Shapley et son application sur la base de données du recensement en Nouvelle-Calédonie.

Ce stage m'a permis d'affiner certaines pistes pour bâtir mon projet d'orientation professionnel et constitue l'aboutissement de mon année scolaire.

Je n'oublie pas non plus mes proches qui m'ont sans cesse soutenu dans l'élaboration de mon projet professionnel.

Résumé

Durant ces 2 mois de stage au sein du centre de recherche en économie et management, j'ai travaillé sur deux projets de recherche que menait Isabelle Lebon, professeur de sciences économiques.

Le premier projet concernait l'étude de données d'expérimentations sur la présidentielle française de 2022. Nous avons étudié et analysé deux bases de données contenant les réponses de 1972 individus sur leur vote suivant 4 modes de scrutins : le scrutin majoritaire, le vote par approbation, le vote par note et le jugement majoritaire.

Dans un deuxième temps, nous avons construit un axe politique endogène à partir des données du vote par approbation. Pour cela, j'ai utilisé un algorithme développé par le précédent stagiaire qui permet d'obtenir un axe de deux manières différentes : en comptant les « trous » puis en gardant le nombre de bulletins sans « trous » ou en additionnant le nombre de « trous » sur l'ensemble des bulletins.

Au final, nous avons conclu que le vote par approbation donnait des résultats similaires au scrutin majoritaire mais que le vote par note et le jugement majoritaire permettaient aux électeurs de « sanctionner » un ou plusieurs candidats en leur attribuant une note négative ou en lui donnant la mention « A rejeter ».

De plus, nous avons obtenu deux axes politique endogènes différents en utilisant les deux algorithmes. Nous sommes arrivés à la conclusion que l'indicateur « au sens strict » donnait des résultats plus aléatoires que celui « au sens large » mais que les deux axes donnaient cependant une représentation fidèle de l'échiquier politique Gauche-Droite tel qu'il était à ce moment-là.

Enfin nous avons étudié la composition des inégalités en Nouvelle-Calédonie grâce à un algorithme de Shapley. Nous avons observé qu'une grande partie des inégalités provenait de la distance « mine la plus proche-habitation » mais que ce pourcentage était plus faible dans les zones urbanisées comme c'est le cas à Nouméa et dans la province Sud plus généralement.

Sommaire

Remerciements.....	1
Résumé.....	2
Introduction.....	5
I- L'environnement économique du stage.....	6
A – Le secteur : La recherche publique en économie.....	6
1. Présentation.....	6
2. Le secteur économique.....	6
B – La structure du laboratoire.....	7
1. L'historique du centre de recherche en économie et management.....	7
2. Le centre de recherche en économie et management aujourd'hui et demain.....	8
3. L'organigramme du laboratoire.....	9
II- Contenu du stage.....	10
A- Analyse de données de sondages pour l'élection présidentielle française de 2022.....	10
1. Contexte.....	10
2. Outils utilisés.....	10
3. Présentation des données.....	11
4. Préparation des données et statistiques descriptives.....	13
5. Résultats selon le mode de scrutins.....	16
B - Construction d'un axe politique endogène à partir du vote par approbation.....	27
1. Contexte et construction de la base de données.....	27
2 . Algorithmes de calcul d'un axe politiques.....	28
3. Limites du mode de calculs.....	32
C – Décomposition des inégalités en Nouvelle-Calédonie.....	32
1. Contexte.....	33
2. Présentation des données.....	33
3. Algorithme de décomposition des inégalités.....	34
4. Résultats.....	36

Conclusion.....	38
Bibliographie.....	39
Annexe.....	40

Introduction

Du 17 avril 2022 au 18 juin 2022, j'ai effectué un stage au sein du centre de recherche en économie et management. Au cours de ce stage au sein du pôle « Politiques Publiques, Vote et Démocratie », j'ai pu m'intéresser à la recherche en économie. Le centre de recherche en économie et management situé à Rennes et Caen, créé en 1990 est spécialisé dans la recherche en sciences économiques et en science de gestion. J'ai été soutenu par Madame Isabelle Lebon, professeur des universités, avec qui j'ai pu me former dans d'excellentes conditions.

Mon stage a consisté dans un premier temps en une analyse de données de panels représentatifs de la population française portant sur l'élection présidentielle de 2022. J'ai ensuite travaillé en collaboration avec Monsieur Frédéric Chantreuil, professeur à l'université de Nouvelle-Calédonie, sur un algorithme de Shapley permettant de calculer des inégalités en fonction de la distance aux mines de Nouvelle-Calédonie.

Cette expérience a été l'opportunité pour moi de percevoir comment un établissement spécialisé dans la recherche en économie se développe et de relever les défis auxquels elle est confrontée, notamment la question de l'analyse des résultats à la présidentielle de 2022 ou du calcul des inégalités.

L'élaboration de ce rapport a pour principale source la pratique journalière des missions qui m'étaient affectées, mise en parallèle avec les enseignements théoriques de ma formation.

Afin de rendre compte de manière fidèle et analytique des 2 mois passés au sein du centre de recherche en économie et management, il apparaît logique de présenter l'environnement économique du stage (I), puis de préciser les différentes missions et tâches que j'ai pu effectuer et les qualités professionnelles et personnelles que j'ai pu en tirer (II).

I- L'environnement économique du stage

A – Le secteur : La recherche publique en économie

1. Présentation

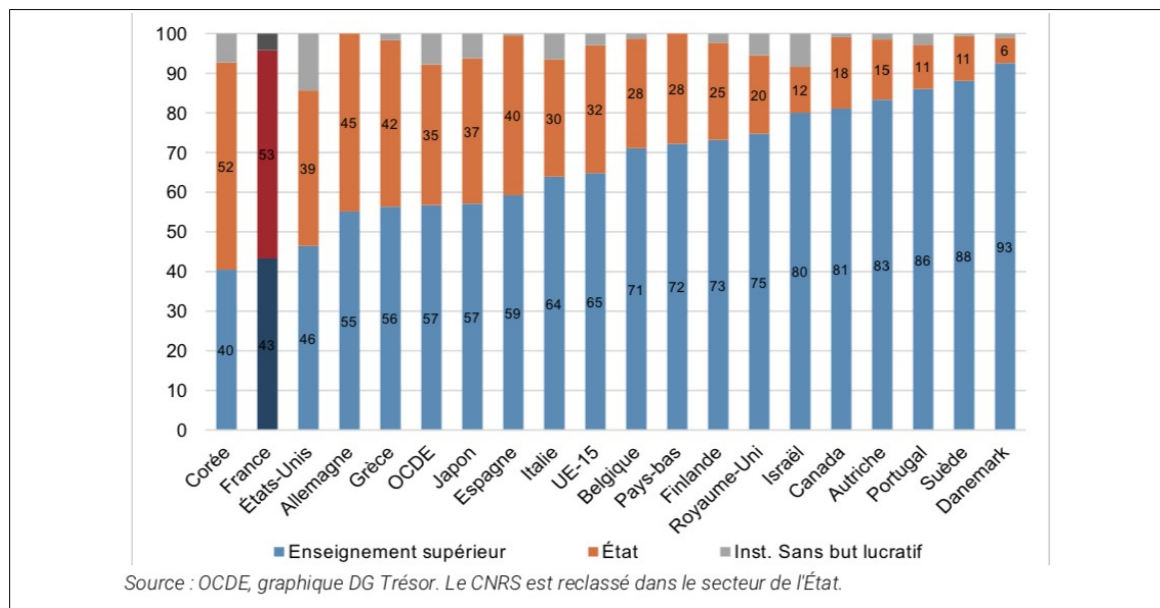
La recherche publique en France est organisée par le ministère de l'Enseignement supérieur et de la recherche. Elle est financée par l'État français et les collectivités territoriales françaises. La recherche publique s'exerce dans trois types d'établissements : les établissements à caractère scientifique, culturel et professionnel (universités, grandes écoles, écoles normales supérieures), les établissements publics à caractère scientifique et technologique (CNRS, INED, INRAE,) et les établissements publics à caractère industriel et commercial (CEA, CNES, ADEME,). Ces établissements emploient différentes personnes contribuant à un travail de recherche comme les doctorants, les maîtres de conférences, les professeurs d'universités. La majorité des chercheurs effectuent des missions d'enseignements au sein d'une université ou d'une grande école. Plus spécifiquement, la recherche en économie est très présente dans les institutions publique comme l'INSEE ou l'INED. Ces institutions produisent des prévisions économiques pour le gouvernement français.

2. Le secteur économique

En 2019, l'investissement dans la recherche publique représentait 2,9% du PIB français, soit la 13^{ème} place du classement des pays de l'OCDE. En 1992, elle occupait la 4^{ème} place de ce classement, soit une chute de 9 places en 27 ans. En 2015, 53% des dépenses intérieure de R&D des administrations était exécuté par l'État, 43% par l'enseignement supérieur et 4% par des instituts à but non lucratif. La France est le pays de l'OCDE où l'État finance le plus la R&D des administrations.

En 2015, les publications scientifiques française représentaient 3,3% du total des publications, ce qui la plaçait au 7^{ème} rang mondial en termes de publication. Les citations françaises représentaient 3,8% des citations mondiales. En 2014, la part des publications en sciences humaines et sociale représentait 7% des publications au niveau national.

Structure de la dépense intérieure de R&D des administrations par type d'exécutant (2015, en %)



B – La structure du laboratoire

1. L'historique du centre de recherche en économie et management

Le centre de recherche en économie et management (CREM) est une unité mixte de recherche (UMR) créée à la fin des années 1990 et regroupant le CNRS, l'université de Rennes 1 et l'université de Caen Normandie. Il fait suite à la fusion progressive des différents UMR de l'ouest de la France en économie et gestion. Depuis 2004, le CREM est le seul UMR de l'ouest de la France dédié aux sciences

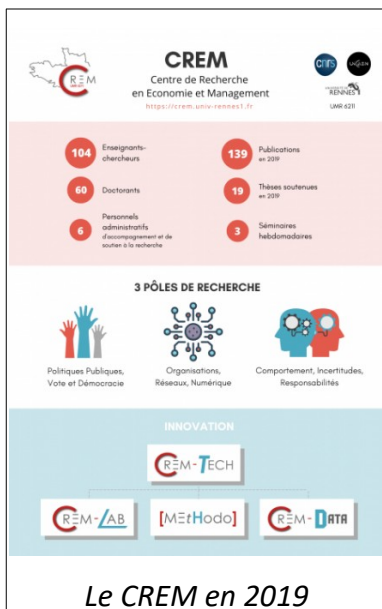


La maison de la recherche en sciences humaines

économiques et de gestion. Le CREM est hébergé au sein de 3 composantes : la faculté des sciences économiques de l'université Rennes 1 (45 chercheurs permanents), l'institut de gestion de Rennes (31 chercheurs permanents) et l'UFR des sciences économiques, de gestion, de géographie et d'aménagement des territoires de l'université de Caen Normandie (25 chercheurs permanents). La Maison de la Recherche en Sciences Humaines (MRSH), une unité d'appui et de recherche CNRS basée à Caen, accueille 6 enseignants-chercheurs du CREM, dont le bureau d'Isabelle Lebon. Depuis 2017, le laboratoire est organisé autour de trois pôles : PPDV (Politiques Publiques, Vote et Démocratie), ORN (Organisation, Réseaux et Numérique) et RRC (Risques, responsabilité et Comportements). Pour ma part, j'ai effectué mon stage au sein du pôle PPDV sur les questions de choix social et de vote.

2. Le centre de recherche en économie et management aujourd'hui et demain

Pour la période 2022-2027, le CREM continuera d'innover dans la recherche: expérimentations au sein du CREM-LAB et du Neuro-Lab grâce aux développements des projets autour de l'économie expérimentale et de la neuroéconomie (Eye



tracking, Imagerie par résonance magnétique fonctionnelle, Électromyographie faciale). Depuis 2011, le CREM développe un partenariat avec le *Condorcet For Politic Center*, une organisation internationale promouvant la recherche dans les domaines de l'économie politique, du choix public, de la théorie de la décision et des finances publiques. De plus, dans le cadre de la transformation digitale du laboratoire, un projet de base de données pour la recherche nommé CREM-DATA a été développé. A terme, le projet permettra d'accéder à l'ensemble des données en lien avec les recherches menées par le CREM. Dès 2017, le laboratoire a expérimenté le télétravail pour ses personnels administratifs situés à Rennes et à Caen notamment

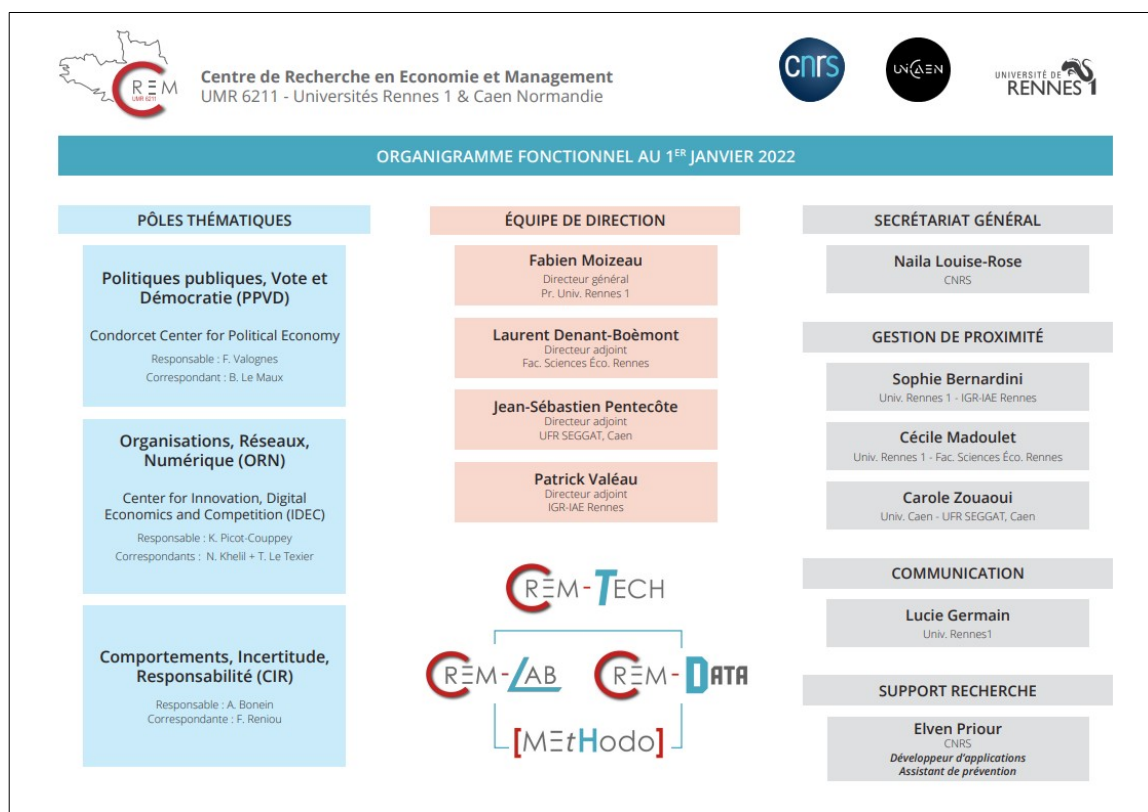
en faisant des dotations de matériels informatiques. L'expérimentation s'est étendue à tout le personnel en 2019 et a continué de se développer dans les années qui ont suivi, ce qui a été particulièrement bénéfique pendant la période de confinement du printemps 2020. Enfin, tous les premiers vendredis de chaque mois est organisé un café-CREM où les personnels se retrouvent autour d'un café et d'un croissant pour discuter de leurs recherches.

3. L'organigramme du laboratoire

La direction du CREM est assurée par un directeur général et des directeurs-adjoints présents sur chaque site:

- Directeur Général : Fabien Moizeau
- Directeur adjoint IGR-IAE : Patrick Valéau
- Directeur adjoint Université Caen Normandie : Jean-Sébastien Pentecôte
- Directeur adjoint Université Rennes 1 : Laurent Denant-Boémont

De plus, pour chaque pôle thématique, il existe un responsable et un correspondant. Enfin, les services supports et administratifs sont organisés autour d'une ou plusieurs responsables.



II- Contenu du stage

Durant ces 2 mois de stage, j'ai travaillé sur un projet principal qui concernait l'analyse de données d'expérimentations pour la présidentielle française de 2022. J'ai aussi travaillé sur un second projet qui visait à étudier la composition des inégalités en Nouvelle-Calédonie. Dans cette section, je présenterai le travail réalisé pour ces 2 projets.

A- Analyse de données de sondages pour l'élection présidentielle française de 2022

1. Contexte

Ce projet s'inscrit dans le cadre de l'opération scientifique « Voter Autrement » mener par des chercheurs en sciences économiques et en informatique de l'université de Caen, Grenoble, Paris (PSE et Lamsade), Saint-Étienne et Strasbourg. Le but de cette opération était d'étudier différents modes de scrutins alternatif au scrutin majoritaire à deux tours utilisé en France pour la présidentielle. Plusieurs types d'expérimentations ont été mis en place : une expérimentation in-situ le 10 avril 2022 dans deux bureaux de vote de Strasbourg, une expérimentation ouverte à tous sur le site internet du projet, des expérimentations dans les laboratoires d'économies partenaires et enfin une expérimentation par sondages soumis par l'institut Dynata entre le 4 et le 10 avril 2022 sur un panel de 2168 individus. Dans le cadre de mon stage, j'ai donc travaillé sur les données provenant de l'expérimentation par sondages.

2. Outils utilisés

Pendant toute la durée de mon stage, j'ai partagé le bureau d'Isabelle Lebon situé à la maison de la recherche en sciences humaines sur le campus 1 de l'université de

Caen. J'avais à ma disposition un ordinateur dont l'unité centrale avait été changée pour pouvoir supporter des calculs lourds. Les logiciels R et Excel étaient pré-installés. Les calculs demandant une plus grande puissance de calcul ont été effectués grâce à l'infrastructure informatique Huma-Num et pilotée par Gérard Chatagnon, ingénieur de recherche à l'université de Saint-Étienne. Les données recueillies étaient rangées dans deux fichiers Excel au format CSV. J'avais aussi accès aux captures d'écran des questions du sondage telles qu'elles étaient présentées aux personnes questionnées.

VOTE PAR APPROBATION

Quels sont les candidat(e)s de l'élection présidentielle de 2022 que vous approuveriez ?
(Les candidats sont présentés dans l'ordre officiel du Conseil Constitutionnel)

Indication : Pour le résultat du vote, lorsqu'un électeur ne se prononce pas sur un candidat, on considère qu'il ne l'approuve pas.

❗ Veuillez sélectionner au moins une réponse

	Non approuvé	Approuvé
Nathalie Arthaud	<input type="radio"/>	<input type="radio"/>
Fabien Roussel	<input type="radio"/>	<input type="radio"/>
Emmanuel Macron	<input type="radio"/>	<input type="radio"/>
Jean Lassalle	<input type="radio"/>	<input type="radio"/>
Marine Le Pen	<input type="radio"/>	<input type="radio"/>
Eric Zemmour	<input type="radio"/>	<input type="radio"/>
	Non approuvé	Approuvé
Jean-Luc Mélenchon	<input type="radio"/>	<input type="radio"/>
Anne Hidalgo	<input type="radio"/>	<input type="radio"/>
Yannick Jadot	<input type="radio"/>	<input type="radio"/>
Valérie Pécresse	<input type="radio"/>	<input type="radio"/>
Philippe Poutou	<input type="radio"/>	<input type="radio"/>
Nicolas Dupont-Aignan	<input type="radio"/>	<input type="radio"/>

Capture d'écran d'une question du sondage sur le vote par approbation

3. Présentation des données

J'avais à ma disposition deux fichiers CSV contenant les réponses des personnes sondées. Les deux fichiers avaient des questions communes et des questions

différentes. Les questions communes portaient sur la description de l'individu (genre, âge, ville, ...) mais aussi sur son inscription sur les listes électorales, sur son vote au scrutin officiel, au vote par approbation et sur son opinion pour les 12 candidats en lice pour l'élection présidentielle française de 2022.

Les personnes ayant répondu au questionnaire A étaient confrontées à des questions concernant le vote par note. Celles ayant répondu au questionnaire B devaient expérimenter le vote par jugement majoritaire.

Voici une description des scrutins testés:

- **Le scrutin majoritaire à deux tours**

C'est le scrutin utilisé actuellement en France pour les élections présidentielles. Il se déroule en deux tours. Au premier tour, les électeurs votent pour un seul candidat ou vote blanc ou nul. Les deux candidats ayant récolté le plus de voix s'affrontent au deuxième tour. Au deuxième tour, les électeurs choisissent un des deux candidats ou votent blancs ou nul. Le candidat qui l'emporte est celui qui récolte le plus de voix. Seul le premier tour a été expérimenté pour ce projet.

- **Le vote par note**

Le vote par note est un scrutin à un tour. Dans notre expérimentation, les électeurs donnent une note à chaque candidat sur une échelle de -1 à 4. Ils peuvent donner la même note à tous les candidats. Le candidat qui l'emporte est celui dont la somme des notes attribuées par chaque électeur est la plus élevée. Dans notre expérimentation, les électeurs avaient le choix de ne donner aucune note à un ou plusieurs candidats. Dans ce cas, une note de -1 était attribuée pour les candidats n'ayant pas de note. Seules les personnes ayant répondu au questionnaire A ont été questionnées sur le vote par note.

- **Le vote par approbation**

Le vote par approbation est lui aussi un scrutin à un tour. On demande aux électeurs si ils approuvent ou désapprouvent chaque candidat. L'électeur peut approuver ou désapprouver plusieurs candidats. Les électeurs avaient aussi le choix de ne pas se prononcer sur plusieurs candidats. Dans ce cas, on considère que l'électeur n'approuve pas le ou les candidats. Le candidat qui l'emporte est celui qui aura récolté le plus d'approbation de la part des électeurs.

- **Le jugement majoritaire**

Testé uniquement sur les personnes ayant répondu au questionnaire B, le jugement majoritaire est un scrutin à un tour. Pour ce scrutin, les électeurs étaient invités à attribuer à chaque candidat une des mentions de la liste suivante : A rejeter, Insuffisant, Passable, Assez Bien, Bien, Très Bien. En cas de non réponse sur un candidat, la mention « A rejeter » lui était attribuée. Les électeurs pouvaient attribuer la même mention à plusieurs candidats. Pour désigner le vainqueur, il faut calculer la mention médiane de chaque candidat. Le candidat ayant la mention médiane « la plus favorable » est élu.

Nous avons aussi demandé aux électeurs d'attribuer un score d'opinion allant de 0 à 20 pour chaque candidat. Les électeurs pouvaient aussi indiquer qu'ils n'avaient pas d'opinion pour un ou plusieurs candidats. Un même score pouvait être attribué à plusieurs candidats. Une absence de réponse était considérée comme « pas d'opinion » .

Au final, 1079 personnes avaient répondu au questionnaire A et 1043 au questionnaire B. Chaque questionnaire comportait 25 questions (voir annexe 1).

4. Préparation des données et statistiques descriptives

Avant d'effectuer une analyse statistique des données, nous nous sommes questionnés sur la prise en compte ou non de l'ensemble des individus. En effet, nous avons à notre disposition le temps réalisé par chaque individu pour finir le questionnaire et nous avons remarqué qu'un certains nombre d'individus avait

terminé le questionnaire très rapidement. Après concertation, nous avons décidé de conserver uniquement les individus ayant répondu au questionnaire en plus de 3 minutes car nous estimions qu'il était impossible de répondre consciencieusement au questionnaire en moins de 3 minutes. Pour ce faire, j'ai ajouté dans les fichiers CSV une colonne **Pondération** égale à 1 si l'individu avait répondu en plus de 3 minutes et à 0 sinon. J'ai ensuite importé les bases dans R puis j'ai créé deux *data frame* nommés **BaseAPondéré** et **BaseBPondéré** ne contenant que les individus pour lesquels la pondération était de 1.

1	Êtes-vous in	Quel âge av	Quel est votr	Quelle est vc	De quelle rée	Quelle est la	La méthode	Si vous votie	Vous allez vc	Vous allez vc	Au cas où vc	Le vote par	Le score obt	A votre avis	Quels
2	Oui	25-34 ans	Femme	Agriculteur, a	Ile-de-France	Agglomération	Oui	Jean-Luc Mélenchon				J'ai bien compris ce mode	Approuvé	Non aj	
3	Oui	35-49 ans	Femme	Employé	Hauts-de-Fra	Moins de 2 0	Oui	Marine Le Pen				J'ai bien compris ce mode	Approuvé	Non aj	
4	Oui	35-49 ans	Femme	Inactif	Occitanie	20 000 à 99	Oui	Yannick Jadot				J'ai bien compris ce mode	Approuvé	Approu	
5	Oui	25-34 ans	Femme	Inactif	Ile-de-France	Agglomération	Je ne sais pas encore				Marine Le Pe	Je voudrais plus d'explicat	Approuvé	Non aj	
6	Oui	50-64 ans	Homme	Cadre et prof	Ile-de-France	20 000 à 99	Oui	Emmanuel Macron				J'ai bien compris ce mode	Approuvé	Non aj	
7	Oui	35-49 ans	Homme	Cadre et prof	Grand Est	Moins de 2 0	Oui	Jean-Luc Mélenchon				J'ai bien compris ce mode	Approuvé	Approu	
8	Oui	50-64 ans	Femme	Employé	Occitanie	2 000 à 19 9	Oui	Nicolas Dupont-Aignan				J'ai bien compris ce mode	Approuvé	Non aj	
9	Oui	65 ans et plu	Homme	Inactif	Ile-de-France	Agglomération	Oui	Eric Zemmour				J'ai bien compris ce mode	Approuvé	Non aj	
10	Oui	25-34 ans	Femme	Profession in	Ile-de-France	Agglomération	Oui	Valérie Pécresse				J'ai bien compris ce mode	Approuvé	Non aj	

Extrait du fichier Base A.csv

Après avoir enlevé les individus ayant répondu en moins de 3 minutes, la base A ne contenait plus que 1002 individus et la base B 970 individus

Pour vérifier que nos échantillons étaient toujours représentatifs de la population française, nous avons effectué des statistiques descriptives sur les caractéristiques des individus dans R avec la commande *table*. Le tableau suivant montre que l'on a gardé un échantillon représentatif de la population française.

		Échantillon A	Échantillon B
Genre	Femmes	52,30 %	52,78 %
	Hommes	47,70 %	47,22 %
Age	18-24 ans	7,68 %	7,53 %
	25-34 ans	14,87 %	12,16 %
	35-49 ans	25,25 %	26,08 %
	50-64 ans	26,15 %	27,32 %

	65 ans et plus	26,05 %	26,91 %
CSP	Agriculteurs, artisans, commerçants, chefs d'entreprise	4,59 %	4,33 %
	Cadres et professions intellectuelles supérieures	11,28 %	11,03 %
	Professions intermédiaires	16,27 %	16,19 %
	Employés	18,06 %	17,94 %
	Ouvriers	9,28 %	9,07 %
	Retraités	27,94 %	29,69 %
	Inactifs	12,57 %	11,75 %
Agglomération	Commune rurale (- de 2000 hab.)	16,07 %	16,91 %
	2000 à 19999 habitants	25,85 %	24,95 %
	20000 à 99999 habitants	17,76 %	19,38 %
	100000 habitants et plus	30,94 %	30,31 %
	Agglomération parisienne	9,38 %	8,45 %

Nous avons alors commencé à étudier le résultat de l'élection selon le mode de scrutin choisis.

5. Résultats selon le mode de scrutins

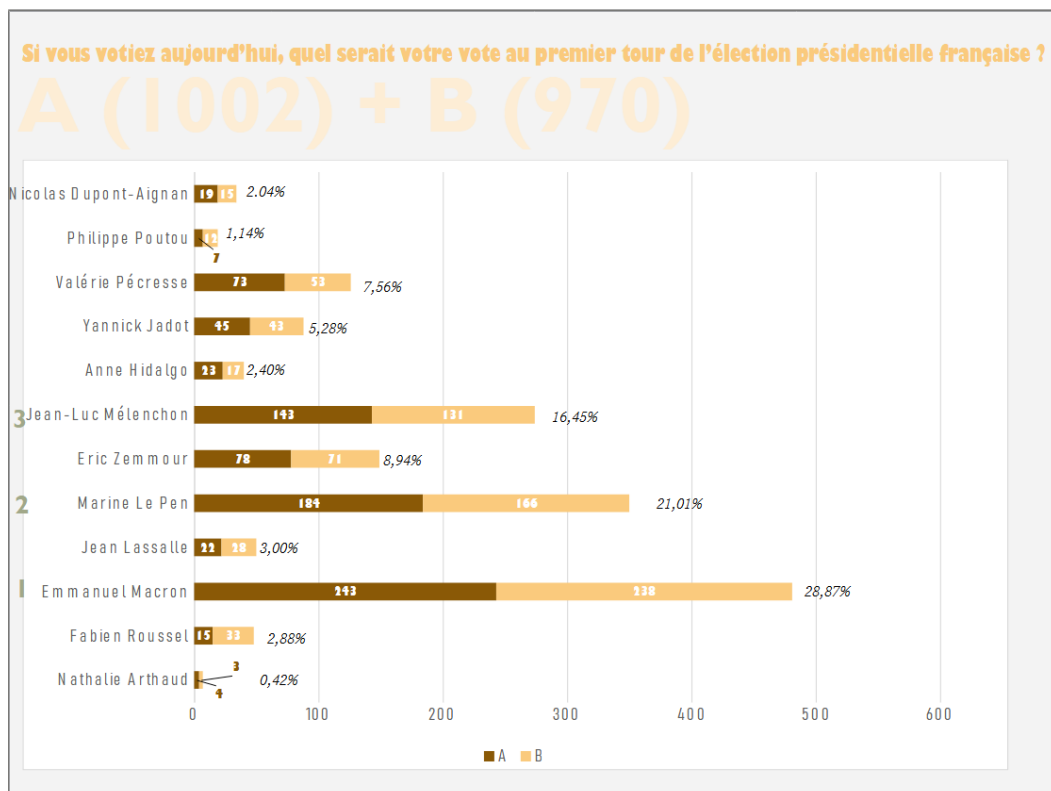
- **Résultats du scrutin majoritaire**

Pour obtenir un classement sur le scrutin majoritaire, j'ai utilisé dans R la fonction *table* sur la base A et la base B en prenant en indice uniquement le numéro de la colonne où la question était posée.

```
1 install.packages("ggplot2")
2 install.packages("xlsx")
3 library("xlsx")
4 library(ggplot2)
5 options(java.parameters = "-Xmx1000m")
6 #On importe la base de données A munie de la colonne pondération ----
7 BaseA <- read.csv("C:/Data/VOTE_DYNATA/Statistiques programme R/Données Pondéré A et B/A_Fusion_Pondéré_2022.04.11_14h24_complétés.csv", header=TRUE, sep=";")
8
9 #On ne garde que les lignes dont la pondération est différente de 0 ----
10 BaseAPondéré <- subset(BaseA, Pondération!= 0)
11 table(BaseAPondéré[,17])
12 write.xlsx(x = BaseAPondéré, file = "BaseAPondéré.xlsx", row.names=FALSE)
13 #Répartition des bulletin par vote officiel déclaré en effectif et pourcentage ----
14 table_vote_officiel <- table(BaseAPondéré[,20])
15 print(table_vote_officiel)
16 prop.table(table_vote_officiel)*100
17 write.xlsx(x = table_vote_officiel, file = "Répartition des bulletin par vote officiel déclaré en effectif et pourcentage.xlsx", row.names=FALSE)
```

Code R pour le scrutin majoritaire

Nous obtenons les résultats suivant en effectifs et pourcentages du nombre de voix sur les bases A et B :



Résultats du vote officielle en effectifs et pourcentages

Pour les deux bases, nous obtenons un classement légèrement différent de celui obtenu à l'issue du premier tour de la présidentielle le 10 Avril 2022, avec une inversion des places entre Nicolas Dupont-Aignan et Anne Hidalgo. Le pourcentage de voix obtenues par Anne Hidalgo est supérieur à celui obtenu le 10 Avril de même que le pourcentage de voix obtenues par Valérie Pécresse est supérieur à celui obtenu pour l'élection officielle alors que Jean-Luc Mélenchon obtient un pourcentage de voix inférieur à celui obtenu le 10 Avril.

- **Résultats du vote par approbation**

Pour obtenir le résultat du vote par approbation, j'ai utilisé la fonction *table* sur les base A et B pour chacune des 12 colonnes contenant les données d'approbation des candidats.

```
68 #Répartition des approbations par candidats ----
69 table(BaseAPondéré[,27])
70 table(BaseAPondéré[,28])
71 table(BaseAPondéré[,29])
72 table(BaseAPondéré[,30])
73 table(BaseAPondéré[,31])
74 table(BaseAPondéré[,32])
75 table(BaseAPondéré[,33])
76 table(BaseAPondéré[,34])
77 table(BaseAPondéré[,35])
78 table(BaseAPondéré[,36])
79 table(BaseAPondéré[,37])
80 table(BaseAPondéré[,38])
```

Code R pour le vote par approbation

Nous avons alors les résultats suivants en pourcentage d'approbations et de désapprobation pour la base A et la base B :

Nombre d'approbations par candidat et pourcentage d'approbations par candidat

	Base A	Base B	Classement
N. Arthaud	156 (15,5688 % par nb bulletins)	127 (13,0927%)	12
F. Roussel	190 (18,9620%)	204 (21,0309%)	8
E. Macron	447 (44,6107%)	424 (43,7113%)	1
J. Lassalle	212 (21,1576%)	214 (22,0618%)	7

M. Le Pen	394 (39,3213%)	356 (36,7010%)	2
E. Zemmour	225 (22,4550%)	220 (22,6804%)	5
JL Mélenchon	325 (32,4351%)	310 (31,9587%)	3
A.Hidalgo	182 (18,1636%)	168 (17,3195%)	9
Y. Jadot	229 (22,8542%)	211 (21,7525%)	6
V. Péresse	260 (25,9481%)	233 (24,0206%)	4
P. Poutou	178 (17,7644%)	162 (16,7010%)	11
N. Dupont-Aignan	182 (18,1636%)	161 (16,5979%)	10

Nous voyons que par rapport au scrutin officiel, seul un candidat, Eric Zemmour, a été désavantagé et passe de la 4ème place à la 5ème place. Valérie Péresse bénéficie elle du vote par approbation puisqu'elle passe de la 5ème à la 4ème place

En étudiant le nombre d'approbations par bulletins, nous pouvons remarquer qu'une grande partie des électeurs agissent comme si il votait au scrutin officiel puisque , sur les deux bases, environ un quart des électeurs n'accorde qu'une approbation sur l'ensemble des 12 candidats.

Approbations par bulletins en effectifs et pourcentages

	Base A	Base B
Bulletins avec 0 approbation (bulletin blanc)	60 (5,9880 %)	63 (6,4948%)
Bulletins avec 1 approbation	249 (24,8502%)	251 (25,8762%)
Bulletins avec 2 approbations	206 (20,5588%)	205 (21,1340%)
Bulletins avec 3 approbations	179 (17,8642%)	145 (14,9484%)
Bulletins avec 4 approbations	106 (10,5788%)	123 (12,6804%)
Bulletins avec 5 approbations	74 (7,3852%)	72 (7,4226%)
Bulletins avec 6 approbations	54 (5,3892%)	46 (4,7422%)
Bulletins avec 7 approbations	24 (2,3952%)	25 (2,5773%)
Bulletins avec 8 approbations	14 (1,3972%)	16 (1,6494%)
Bulletins avec 9 approbations	8 (0,7984%)	5 (0,5154%)
Bulletins avec 10 approbations	10 (0,9980%)	5 (0,5154%)
Bulletins avec 11 approbations	4 (0,3992%)	0 (0%)
Bulletins avec 12 approbations	14 (1,3972%)	14 (1,4432%)
TOTAL	1002	970

Ceci explique pourquoi, pour cette expérimentation, nous obtenons à peu près le même classement que pour l'expérimentation du scrutin majoritaire.

- **Résultats du vote par note**

Pour obtenir le classement avec le vote par note, j'ai d'abord créé dans R un *data frame* pour chaque candidat qui était composé de la colonne des réponses à la question du vote par note pour l'ensemble des électeurs, en remplaçant les données vides par la note de -1. J'ai ensuite utilisé la fonction *sum* sur chaque *data frame* pour obtenir un score qui était égal à la somme de l'ensemble des notes obtenues par candidat.

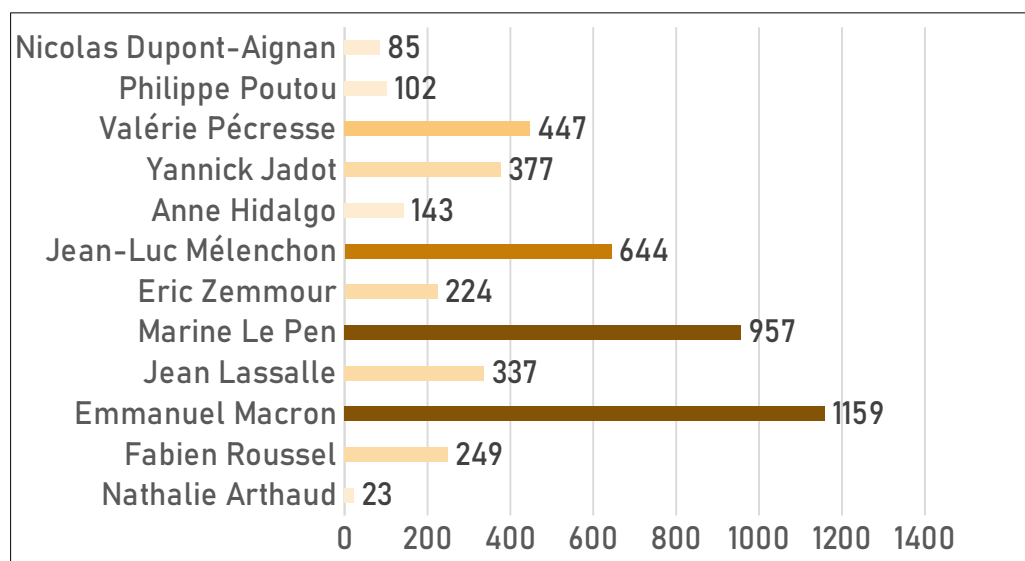
```

136 - #Répartition du vote par note en effectif et note moyenne ----
137 BaseAPondéréNoteNA <- replace(BaseAPondéré[,42], BaseAPondéré[,42]== "", -1)
138 BaseAPondéréNoteFR <- replace(BaseAPondéré[,43], BaseAPondéré[,43]== "", -1)
139 BaseAPondéréNoteEM <- replace(BaseAPondéré[,44], BaseAPondéré[,44]== "", -1)
140 BaseAPondéréNoteJL <- replace(BaseAPondéré[,45], BaseAPondéré[,45]== "", -1)
141 BaseAPondéréNoteMLP <- replace(BaseAPondéré[,46], BaseAPondéré[,46]== "", -1)
142 BaseAPondéréNoteEZ <- replace(BaseAPondéré[,47], BaseAPondéré[,47]== "", -1)
143 BaseAPondéréNoteJLM <- replace(BaseAPondéré[,48], BaseAPondéré[,48]== "", -1)
144 BaseAPondéréNoteAH <- replace(BaseAPondéré[,49], BaseAPondéré[,49]== "", -1)
145 BaseAPondéréNoteYJ <- replace(BaseAPondéré[,50], BaseAPondéré[,50]== "", -1)
146 BaseAPondéréNoteVP <- replace(BaseAPondéré[,51], BaseAPondéré[,51]== "", -1)
147 BaseAPondéréNotePP <- replace(BaseAPondéré[,52], BaseAPondéré[,52]== "", -1)
148 BaseAPondéréNoteNDA <- replace(BaseAPondéré[,53], BaseAPondéré[,53]== "", -1)
149 #Nathalie Arthaud
150 sum(BaseAPondéréNoteNA)
151 sum(BaseAPondéréNoteNA)/1002
152 #Fabien Roussel
153 sum(BaseAPondéréNoteFR)
154 sum(BaseAPondéréNoteFR)/1002
155 #Emmanuel Macron
156 sum(BaseAPondéréNoteEM)
157 sum(BaseAPondéréNoteEM)/1002
158 #Jean Lassalle
159 sum(BaseAPondéréNoteJL)
160 sum(BaseAPondéréNoteJL)/1002
161 #Marine Le Pen
162 sum(BaseAPondéréNoteMLP)
163 sum(BaseAPondéréNoteMLP)/1002
164 #Eric Zemmour
165 sum(BaseAPondéréNoteEZ)
166 sum(BaseAPondéréNoteEZ)/1002
167 #Jean-Luc Mélenchon
168 sum(BaseAPondéréNoteJLM)
169 sum(BaseAPondéréNoteJLM)/1002
170 #Anne Hidalgo
171 sum(BaseAPondéréNoteAH)
172 sum(BaseAPondéréNoteAH)/1002
173 #Yannick Jadot
174 sum(BaseAPondéréNoteYJ)
175 sum(BaseAPondéréNoteYJ)/1002
176 #Valérie Pécresse
177 sum(BaseAPondéréNoteVP)
178 sum(BaseAPondéréNoteVP)/1002
179 #Philippe Poutou
180 sum(BaseAPondéréNotePP)
181 sum(BaseAPondéréNotePP)/1002
182 #Nicolas Dupont-Aignan
183 sum(BaseAPondéréNoteNDA)
184 sum(BaseAPondéréNoteNDA)/1002

```

Code R pour le vote par note

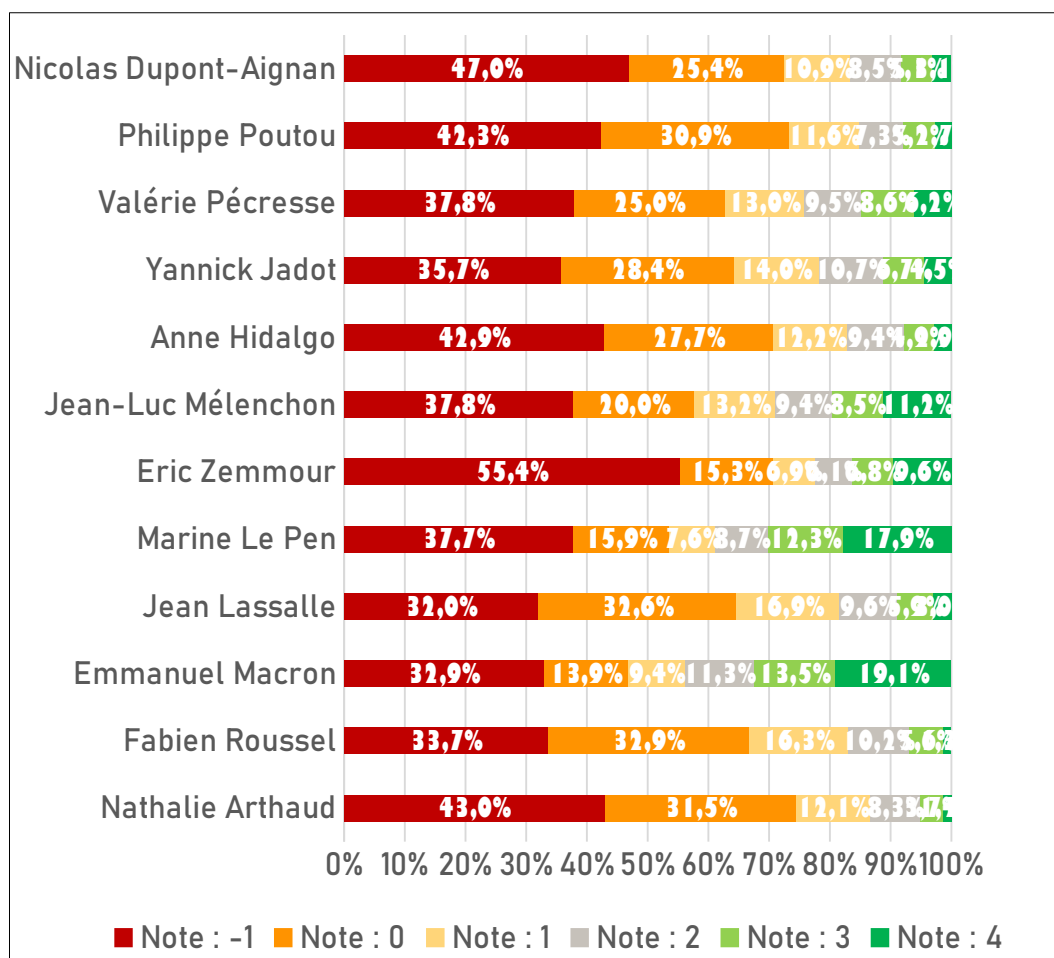
Nous obtenons alors les scores suivants :



Scores du vote par note par candidat

Nous observons plusieurs changements par rapport aux deux scrutins présentés précédemment. Le trio de tête reste le même, mais Eric Zemmour perd 4 places et se classe 8ème. Ce qui profite à Valérie Pécresse, Yannick Jadot, Jean Lassalle et Fabien Roussel qui gagnent une place chacun. Nicolas Dupont-Aignan et Philippe Poutou échangent leur place, devenant respectivement 11ème et 10ème.

Nous avons alors mené une analyse plus poussée en étudiant la répartition des notes par candidat.



Répartition des notes par candidat en pourcentage

Nous observons alors une surreprésentation de la note -1 parmi l'ensemble des notes données aux candidats. En effet, les électeurs ont tendance à utiliser -1 pour sanctionner le ou les candidats qu'ils ne veulent pas voir gagner. Par rapport aux autres modes de scrutin, le vote par note expérimenté ici permet à l'électeur, en plus de désigner son candidat préféré, d'éliminer le candidat qu'il ne veut pas voir gagner.

- **Résultats du jugement majoritaire**

Pour le jugement majoritaire j'ai créé des *data frame* comme pour le vote par note, en remplaçant les données manquantes par « A rejeter ». Puis j'ai utilisé la fonction *table* sur chaque *data frame* pour avoir la répartition des mentions par candidat.

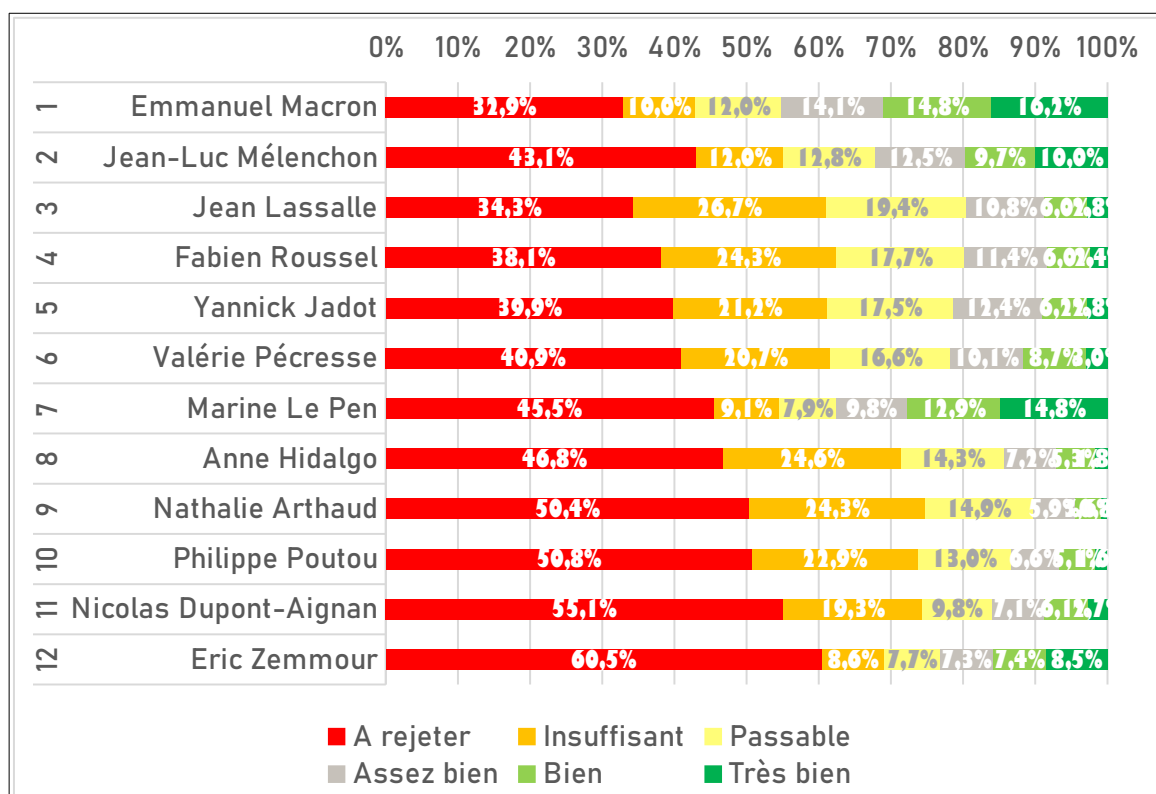
```

118 #Répartition des mention du vote par jugement majoritaire par candidats ----
119 BaseBPondéréMajoritaireNA <- replace(BaseBPondéré[,42],BaseBPondéré[,42]=="","A rejeter")
120 BaseBPondéréMajoritaireFR <- replace(BaseBPondéré[,43],BaseBPondéré[,43]=="","A rejeter")
121 BaseBPondéréMajoritaireEM <- replace(BaseBPondéré[,44],BaseBPondéré[,44]=="","A rejeter")
122 BaseBPondéréMajoritaireJL <- replace(BaseBPondéré[,45],BaseBPondéré[,45]=="","A rejeter")
123 BaseBPondéréMajoritaireMLP <- replace(BaseBPondéré[,46],BaseBPondéré[,46]=="","A rejeter")
124 BaseBPondéréMajoritaireEZ <- replace(BaseBPondéré[,47],BaseBPondéré[,47]=="","A rejeter")
125 BaseBPondéréMajoritaireJLM <- replace(BaseBPondéré[,48],BaseBPondéré[,48]=="","A rejeter")
126 BaseBPondéréMajoritaireAH <- replace(BaseBPondéré[,49],BaseBPondéré[,49]=="","A rejeter")
127 BaseBPondéréMajoritaireYJ <- replace(BaseBPondéré[,50],BaseBPondéré[,50]=="","A rejeter")
128 BaseBPondéréMajoritaireVP <- replace(BaseBPondéré[,51],BaseBPondéré[,51]=="","A rejeter")
129 BaseBPondéréMajoritairePP <- replace(BaseBPondéré[,52],BaseBPondéré[,52]=="","A rejeter")
130 BaseBPondéréMajoritaireNDA <- replace(BaseBPondéré[,53],BaseBPondéré[,53]=="","A rejeter")
131 table_majoritaire_NA <- table(BaseBPondéréMajoritaireNA)
132 table_majoritaire_FR <- table(BaseBPondéréMajoritaireFR)
133 table_majoritaire_EM <- table(BaseBPondéréMajoritaireEM)
134 table_majoritaire_JL <- table(BaseBPondéréMajoritaireJL)
135 table_majoritaire_MLP <- table(BaseBPondéréMajoritaireMLP)
136 table_majoritaire_EZ <- table(BaseBPondéréMajoritaireEZ)
137 table_majoritaire_JLM <- table(BaseBPondéréMajoritaireJLM)
138 table_majoritaire_AH <- table(BaseBPondéréMajoritaireAH)
139 table_majoritaire_YJ <- table(BaseBPondéréMajoritaireYJ)
140 table_majoritaire_VP <- table(BaseBPondéréMajoritaireVP)
141 table_majoritaire_PP <- table(BaseBPondéréMajoritairePP)
142 table_majoritaire_NDA <- table(BaseBPondéréMajoritaireNDA)
143 #Nathalie Arthaud
144 print(table_majoritaire_NA)
145 #Fabien Roussel
146 print(table_majoritaire_FR)
147 #Emmanuel Macron
148 print(table_majoritaire_EM)
149 #Jean Lassalle
150 print(table_majoritaire_JL)
151 #Marine Le Pen
152 print(table_majoritaire_MLP)
153 #Eric Zemmour
154 print(table_majoritaire_EZ)
155 #Jean-Luc Mélenchon
156 print(table_majoritaire_JLM)
157 #Anne Hidalgo
158 print(table_majoritaire_AH)
159 #Yannick Jadot
160 print(table_majoritaire_YJ)
161 #Valérie Pécresse
162 print(table_majoritaire_VP)
163 #Philippe Poutou
164 print(table_majoritaire_PP)
165 #Nicolas Dupont-Aignan
166 print(table_majoritaire_NDA)

```

Code R pour le jugement majoritaire

Nous avons alors obtenu la répartition suivante des mentions par candidat et leur classement :



Répartition des mentions en pourcentage du total par candidat

Ce qui nous donne la mention médiane par candidat :

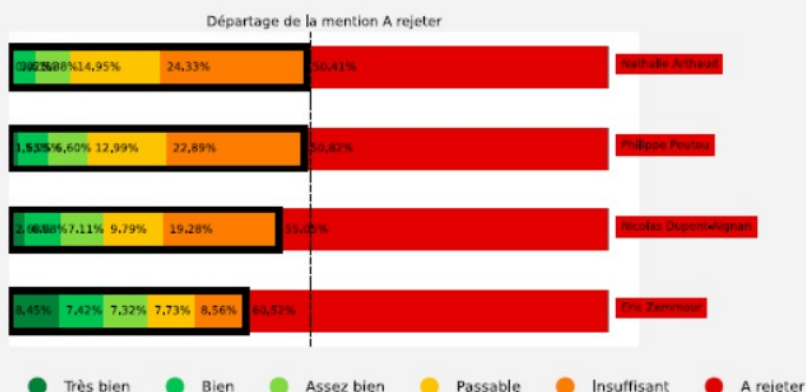
Mention médiane par candidat

	Base B
N. Arthaud	A rejeter
F. Roussel	Insuffisant
E. Macron	Passable
J. Lassalle	Insuffisant
M. Le Pen	Insuffisant
E. Zemmour	A rejeter
JL Mélenchon	Insuffisant
A.Hidalgo	Insuffisant
Y. Jadot	Insuffisant
V. Pécresse	Insuffisant
P. Poutou	A rejeter
N. Dupont-Aignan	A rejeter

Ce résultat nous donne un vainqueur, Emmanuel Macron, puisque c'est le seul candidat à avoir la mention médiane « Passable ». Nous avons ensuite 7 égalités avec la mention médiane « Insuffisant » puis 4 égalités avec la mention médiane « A rejeter ».

Pour la mention « A rejeter », le départage s'effectue comme suit :

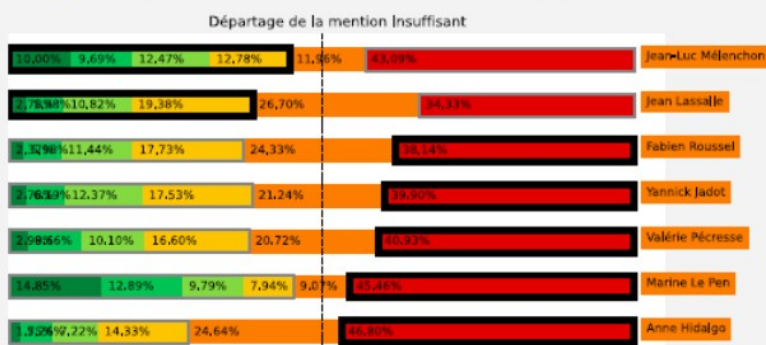
Méthode de départage de la mention "A rejeter" avec le jugement majoritaire:
Le plus grand bloc d'électeurs en désaccord avec la mention majoritaire décide.



- Nicolas Dupont-Aignan a obtenu la mention majoritaire "A rejeter". 44,95% des répondants auraient préféré lui accorder une mention strictement meilleure, et 0,00% des répondants une mention strictement moins bonne que "A rejeter".
- Nathalie Arthaud a également obtenu la mention majoritaire "A rejeter". 49,59% des répondants auraient préféré lui accorder une meilleure mention, et 0,00% des répondants une mention strictement moins bonne que "A rejeter".
- Le plus grand bloc de répondants en désaccord avec la mention majoritaire est composé par ceux qui auraient préféré leur accorder une meilleure mention. Ce bloc est plus important pour Nathalie Arthaud (49,59%) que pour Nicolas Dupont-Aignan (44,95%). Nathalie Arthaud est donc classée avant Nicolas Dupont-Aignan.

Pour la mention « Insuffisant », le départage s'effectue comme suit :

Méthode de départage de la mention "Insuffisant" avec le jugement majoritaire:
Le plus grand bloc d'électeurs en désaccord avec la mention majoritaire décide.



- Jean-Luc Mélenchon a obtenu la mention majoritaire "Insuffisant". 43,09% des répondants ont attribué à Jean-Luc Mélenchon la mention "A rejeter", moins bonne que sa mention majoritaire ; 44,95% (10%+9,69%+12,47%+12,78%) lui ont attribué une meilleure mention que sa mention majoritaire. Le plus grand bloc d'électeurs en désaccord avec la mention majoritaire est donc composé par ceux qui auraient préféré lui accorder une meilleure mention. La mention de départage de Jean-Luc Mélenchon est donc meilleure que insuffisant, c'est la mention "Passable".
- Sept candidats obtiennent la mention majoritaire "Insuffisant". Six candidats obtiennent la mention de départage "A rejeter". Marine Le Pen compte moins de répondants qui auraient préféré lui attribuer une moins bonne mention que "Insuffisant" (45,46%) que Anne Hidalgo (46,80%). Elle est donc classée devant Anne Hidalgo.

Comme pour le vote par note, le jugement majoritaire permet aux électeurs d'exprimer leurs rejets de certains candidats. Par rapport au scrutin officiel : Eric Zemmour perd 8 places et se place à la dernière place du classement. A l'inverse, Jean Lassalle gagne 4 places et monte à la troisième marche du podium. Marine Le Pen perd 5 places et se place à la 7ème place alors que Fabien Roussel gagne. Mais ce mode de scrutin comporte deux problèmes majeurs.

Premièrement, la majorité des électeurs ne comprend pas la notion de médiane et ne peut donc pas départager elle-même les candidats avec le jugement majoritaire. Dans l'expérimentation, une question portait sur la compréhension du jugement majoritaire.

« Sachant que le(la) candidat(e) qui a la meilleure mention est élu(e) » : Réponses correctes surlignées. Le complémentaire c'est « je ne sais pas »











	Base B 180+	
	VRAI	FAUX
Correspond à la mention moyenne donnée à ce candidat par l'ensemble des électeurs	514 (52,9896%)	172 (17,7319%)
Correspond à la mention moyenne donnée à ce candidat par la moitié plus un des électeurs qui lui sont le plus favorables.	350 (36,0824%)	275 (28,3505%)
Correspond à la mention donnée par l'électeur qui partage en deux l'ensemble des électeurs classés du plus favorable au plus défavorable.	284 (29,2783%)	259 (26,7010%)

Sur 970 personnes interrogées, environ 30 % comprennent le mode de départage. La méconnaissance du mode de départage du jugement majoritaire par les électeurs introduit un biais dans les réponses qui peut fausser le classement final.

Le deuxième problème du scrutin majoritaire est le mode de départage en cas d'égalité. En effet, avec 5 mentions possibles, il suffit qu'il y ait plus de 5 candidats en lice pour qu'il y ait obligatoirement égalité de mention médiane entre minimum 2 candidats. Les égalités demandent un nouveau départage. Ici, nous avons utilisé un mode de départage nous donnant un podium Macron-Mélenchon-Lassalle. Mais nous avons aussi testé un autre mode de départage consistant à retirer la mention la plus basse à chaque candidat à égalité jusqu'à qu'un ou plusieurs candidats changent de mention médiane et de recommencer jusqu'à ce que tous les candidats soient classés. Avec ce mode de départage, nous avons obtenu un classement Macron-Lassalle-Mélenchon.

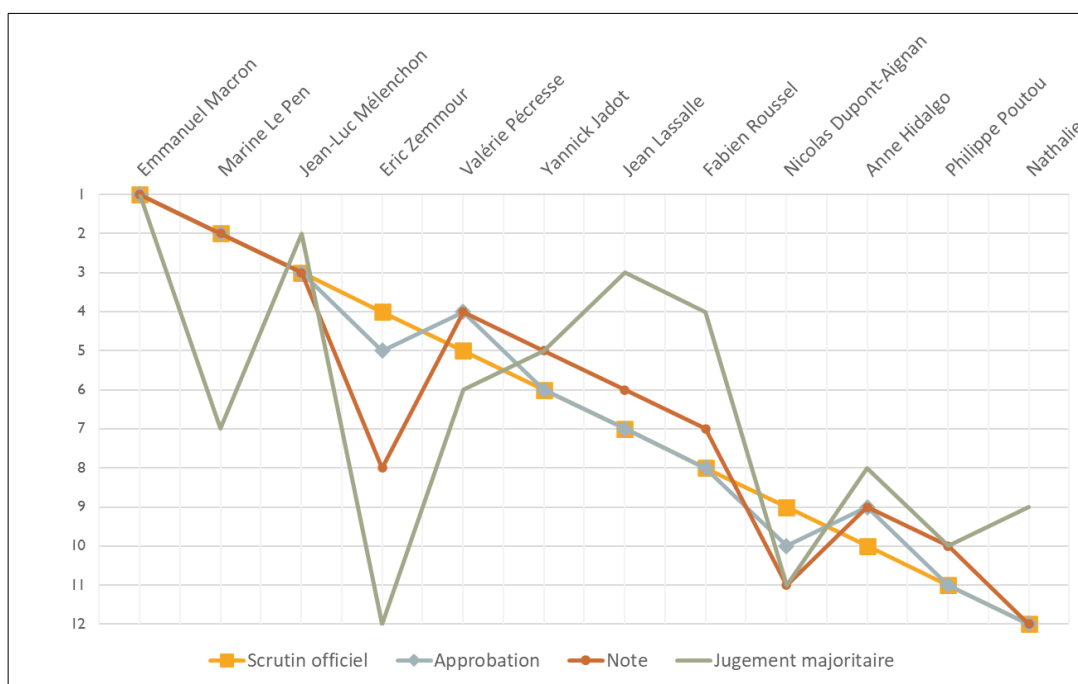
En conclusion, nous avons vu que le vote au scrutin majoritaire et le vote par approbation donnent des résultats assez similaires pour cette expérimentation

alors que le vote par note et le jugement majoritaire permettent de sanctionner un ou plusieurs candidat avec la note -1 et la mention « A rejeter ».

	Premier	Second	Troisième
Uninominal à deux tours			
Approbation			
Note			
Jugement majoritaire			

Comparaison du premier, deuxième et troisième candidat selon le mode de scrutin

En comparant les résultats des 4 modes de scrutins testés, on remarque que le jugement majoritaire est celui qui s'écarte le plus des trois autres, notamment à cause du manque de compréhension pour ce mode de scrutin.



Comparaison du classement des 12 candidats selon le mode de scrutin

B - Construction d'un axe politique endogène à partir du vote par approbation

1. Contexte et construction de la base de données

Après avoir analysé les données de l'expérimentation nous permettant d'obtenir des résultats pour 4 modes de scrutin différents, nous avons à construire un axe endogène représentant l'échiquier politique en 2022. Pour ce faire nous avons utilisé les données du vote par approbation des deux bases A et B. J'ai commencé par transformer ces données de telle sorte que « approuvé » soit remplacé par 1 et « non approuvé » ou une donnée manquante soit remplacée par 0. J'ai ensuite utilisé la fonction *as.numeric* pour transformer les données en variables numériques. Pour finir, j'ai créé un *data frame* contenant les colonnes d'approbations des 12 candidats plus une treizième colonne représentant le nombre d'approbations données par chaque individu, soit ici la somme des « 1 » pour chaque ligne.

```
25 BaseABPondéréBulletins1 <- replace(BaseABPondéré, BaseABPondéré=="Approuvé", 1)
26 BaseABPondéréBulletins2 <- replace(BaseABPondéréBulletins1, BaseABPondéréBulletins1=="Non approuvé", 0)
27 BaseABPondéréBulletins <- replace(BaseABPondéréBulletins2, BaseABPondéréBulletins2=="", 0)
28 BaseABPondéréBulletins[,1] = as.numeric(unlist(BaseABPondéréBulletins[,1]))
29 BaseABPondéréBulletins[,2] = as.numeric(unlist(BaseABPondéréBulletins[,2]))
30 BaseABPondéréBulletins[,3] = as.numeric(unlist(BaseABPondéréBulletins[,3]))
31 BaseABPondéréBulletins[,4] = as.numeric(unlist(BaseABPondéréBulletins[,4]))
32 BaseABPondéréBulletins[,5] = as.numeric(unlist(BaseABPondéréBulletins[,5]))
33 BaseABPondéréBulletins[,6] = as.numeric(unlist(BaseABPondéréBulletins[,6]))
34 BaseABPondéréBulletins[,7] = as.numeric(unlist(BaseABPondéréBulletins[,7]))
35 BaseABPondéréBulletins[,8] = as.numeric(unlist(BaseABPondéréBulletins[,8]))
36 BaseABPondéréBulletins[,9] = as.numeric(unlist(BaseABPondéréBulletins[,9]))
37 BaseABPondéréBulletins[,10] = as.numeric(unlist(BaseABPondéréBulletins[,10]))
38 BaseABPondéréBulletins[,11] = as.numeric(unlist(BaseABPondéréBulletins[,11]))
39 BaseABPondéréBulletins[,12] = as.numeric(unlist(BaseABPondéréBulletins[,12]))
40
41
42
43 base = cbind(BaseABPondéréBulletins[,2], BaseABPondéréBulletins[,3], BaseABPondéréBulletins[,4], BaseABPondéréBulletins[,5],
44             BaseABPondéréBulletins[,6], BaseABPondéréBulletins[,7], BaseABPondéréBulletins[,8], BaseABPondéréBulletins[,9],
45             BaseABPondéréBulletins[,10], BaseABPondéréBulletins[,11], BaseABPondéréBulletins[,12],
46             (BaseABPondéréBulletins[,2]+BaseABPondéréBulletins[,3]+BaseABPondéréBulletins[,4]+BaseABPondéréBulletins[,5]+
47              BaseABPondéréBulletins[,6]+BaseABPondéréBulletins[,7]+BaseABPondéréBulletins[,8]+BaseABPondéréBulletins[,9]+
48              BaseABPondéréBulletins[,10]+BaseABPondéréBulletins[,11]+BaseABPondéréBulletins[,12]))
49
50 write.xlsx(x = base, file = "BaseAB_Axe.xlsx", row.names=FALSE)
```

Programme R pour la construction de la base de données approbation

J'ai ensuite rajouté une dernière colonne **Pondération** sur Excel avec 1 pour l'ensemble des lignes au cas où on souhaiterait corriger les données.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	NAT	FR	EM	JL	MLP	EZ	JLM	AH	YJ	VP	PP	NDA	APPROB	POND
2	1	1	0	0	0	0	0	0	1	0	1	0	4	1
3	0	0	1	0	1	0	1	0	0	1	0	0	4	1
4	0	0	1	0	0	0	0	0	0	1	0	0	2	1
5	1	1	0	0	0	0	1	1	1	0	1	0	6	1
6	0	0	0	0	1	0	0	0	0	1	0	1	3	1
7	0	1	0	0	1	0	0	0	0	0	0	0	2	1
8	1	1	0	1	1	0	1	0	0	1	0	1	7	1
9	0	0	0	0	1	0	0	0	0	0	0	1	2	1
10	0	0	1	1	1	1	0	0	0	0	0	0	4	1
11	0	0	1	0	0	0	0	0	0	1	0	0	2	1
12	0	0	0	0	1	0	0	0	0	0	1	0	2	1
13	0	0	0	0	0	0	1	1	0	0	1	0	3	1
14	1	1	0	0	0	0	1	1	1	0	1	0	6	1
15	0	0	1	1	1	1	0	0	0	0	0	1	5	1
16	0	0	1	0	1	0	1	1	0	1	0	0	5	1
17	0	0	0	0	1	0	1	0	0	0	0	0	2	1
18	0	0	1	0	0	0	0	0	1	0	0	0	2	1
19	0	1	0	1	1	0	0	0	0	0	0	0	3	1
20	0	0	1	0	1	0	0	0	1	0	0	0	3	1
21	1	1	0	1	1	0	0	1	1	1	0	1	8	1
22	0	0	0	0	1	1	1	0	0	0	0	1	4	1

Extrait de la base de données approbation

2 . Algorithmes de calcul d'un axe politiques

Pour obtenir un axe politique endogène, nous avons utilisé deux algorithmes mis au point par madame Lebon. Le premier utilise un indicateur nommé « unimodalité au sens large » et le second un indicateur « unimodalité au sens strict ». Ces indicateurs partent du principe que si un électeur approuve plusieurs candidats, alors c'est qu'il estime que ces candidats sont proches de ses idées, ce qui rend proche ces candidats les uns des autres dans l'esprit de l'électeur. A partir d'un axe donné, les approbations données uniquement à des candidats voisins sur cet axe sont strictement unimodaux.

J'ai utilisé les algorithmes construits par Alexis Léon, stagiaire ayant aussi travaillé sur les axes pour les élections 2017, en les adaptant au contexte de la présidentielle de 2022. Voici leur fonctionnement :

- **Unimodalité au sens large**

On commence par compter le nombre de trous pour chaque bulletin, c'est à dire le nombre de candidats non approuvés ayant deux voisins approuvés.

On additionne ensuite l'ensemble du nombre de trous obtenu à chaque bulletin pour obtenir le score avec l'unimodalité au sens large pour l'axe défini.

```
52 > ponderation=function(i){  
53   a=which(donnes[i,1:12]== 1,arr.ind=TRUE)  
54   b1=list(a[,2])  
55   e2=max(b1[[1]])  
56   e1=min(b1[[1]])  
57   return(as.numeric((e2-e1-donnes$APPROB[i]+1)*(donnes$POND[i])))  
58 }
```

Code R de la fonction calculant le score d'un axe avec l'unimodalité au sens large

- **Unimodalité au sens strict**

Pour obtenir le score d'un axe avec l'unimodalité au sens strict, on compte le nombre de trous par bulletins, puis on garde le nombre de bulletins qui ne comportent aucun trou. Ce nombre de bulletins est le score obtenu avec l'unimodalité au sens strict.

```
60 > ponderation2=function(i){  
61   a=which(donnes[i,1:12]== 1,arr.ind=TRUE)  
62   b1=list(a[,2])  
63   e2=max(b1[[1]])  
64   e1=min(b1[[1]])  
65   donnes$pp <- as.numeric((e2-e1-donnes$APPROB[i]+1))  
66   if (donnes$pp[i]==0)  
67     {return(as.numeric(donnes$POND[i]))}  
68 }
```

Code R de la fonction calculant le score d'un axe avec l'unimodalité au sens strict

En théorie, les deux algorithmes doivent tourner sur un axe composé de 12 candidats. Avec 12 ! permutations, il existe donc 479 001 600 axes possibles. Multipliés par le nombres de bulletins dans la base approbation construite précédemment, cela donne 1972*479001600 calculs à réaliser. Avec l'ordinateur à

notre disposition, nous avons calculé qu'il fallait environ 500 jours pour que l'algorithme puisse effectuer tous les calculs et nous fournisse un résultat.

Nous avons donc eu le soutien de Gérald Chatagnon, ingénieur de recherche, qui avait accès à l'infrastructure informatique Huma-Num. Cet outil nous a permis d'avoir accès à une plus grande puissance de calcul mais aussi de pouvoir paralléliser les programmes afin que la machine puisse faire tourner plusieurs axes prédéfinis en même temps.

Nous avons donc décidé dans un premier temps de faire tourner l'algorithme sur un axe où seulement 6 candidats permutés, les autres candidats étant bloqués. Ceci nous a permis de réduire le temps de calculs à environ une heure. Nous avons alors décidé de bloquer les candidats habituellement classés à gauche (Arthaud, Poutou, Roussel, Mélenchon, Hidalgo, Jadot) au centre (Lassalle, Macron) ou à droite (Pécresse, Le Pen, Zemmour, Dupont-Aignan).

```
1-#Etape 1: on importe les packages nécessaire-sélectionner + ctrl-entrée----
2 setInternet2(use=TRUE)
3 install.packages("gtools")
4 install.packages("xlsx")
5 install.packages("gmodels")
6
7 library("gtools")
8 require(xlsx)
9 library(gmodels)
10
11-#Etape 2: on choisit le chemin dans lequel se trouve la base et on inscrit le nom de la base dans la 2ème ligne----
12 setwd("c:/data/VOTE_OYNATA/statistiques programme R/bases")
13 donnesource=read.csv2("baseAB_Axe_18_34.csv",header=TRUE)
14
15
16-#Etape 3: on entre le nom de tout les candidats----
17 listecandidats=c("NAT","FR","EM","JL","MLP","EZ","JLM","AH","V3","VP","PP","NDA")
18
19
20-#Etape 4: on crée une base regroupant les permutations sur les candidats choisis. n et r représente le nombre de candidat sur lesquels on souhaite faire----
21 #la permutation (ici 6 candidats permutés) (tousjours n=). Ensuite, on change "v=Listecandidats[c()]" pour choisir quels candidats on souhaite permuter
22 #par exemple ici, on permute le candidat en 3ème position dans "listecandidats" qui a été définis au-dessus
23 #ici, le candidat en 3ème position est "EM". Le 10ème est "VP". Le 8ème est "AH". Ainsi de suite jusqu'à avoir tout les candidats que l'on souhaite permuter.
24 base=as.data.frame(permutations(n=6,r=6,v=Listecandidats[c(1,9,8,7,11,2)]))
25
26-#Etape 5: on insère les candidats que l'on souhaite bloqué. Ici, on fait permuté 6 candidats. On ajoute donc une colonne "v7" qui contient le nom du candidat que l'on souhaite ne pas faire bouger----
27 #on recommence pour "v8" qui sera le 2ème candidats que l'on souhaite pas permuté. On continue jusqu'à avoir "v12" puisque l'on a au total 12 candidats.
28 base$V7=rep("EM",120)
29 base$V8=rep("VP",120)
30 base$V9=rep("JL",120)
31 base$V10=rep("MLP",120)
32 base$V11=rep("JL",120)
33 base$V12=rep("NDA",120)
34
35-#Etape 6: on place les candidats dans l'ordre que l'on souhaite. Ici, on place la colonne "v7", celle de Fabien Roussel, en première position.----
36 #v8, celle de Nathalie Arthaud, en deuxième position, et "v9" en troisième position et "v10" en 4ème.
37 #on a bloqué nos 4 candidats à gauche. Nous passons maintenant aux 6 candidats du milieu qui permutent.
38 #ces candidats se trouvent de la colonne "v1" à "v6"
39 #pour finir, on place nos 2 candidats à droite dans l'ordre JL (v11) et NDA (v12)
40 base=base[,c(which(colnames(base))=="V1"),which(colnames(base))=="V2"),which(colnames(base))=="V3"),which(colnames(base))=="V4"),which(colnames(base))=="V5"),which(colnames(base))=="V6"),which(colnames(base))=="V7"),which(c(
```

Code R définissant un axe avec 6 permutations et 6 candidats bloqués

Nous obtenions en sortie une base de données contenant les différents axes avec leurs scores en unimodalité au sens large et au sens strict. Pour obtenir le meilleur axe au sens large, on retient celui qui donne le plus petit score et qui donc minimise le nombre de trous par bulletin. Pour le meilleur axe au sens strict, on prend celui qui

à le plus grand score, c'est à dire celui avec le plus grand nombre de bulletins sans trou.

Nous avons aussi lancé en parallèle plusieurs programmes sur Huma-Num avec un axe composé de 9 permutations et de 3 candidats bloqués. Les résultats de ces programmes nous permettaient d'obtenir un résultat plus précis que celui que l'on obtenait avec seulement 6 permutations.

```

10 setwd("C:/Data/VOTE_DYNATA/Statistiques programme R/Bases")
11 donneeSource=read.csv2("BaseAB_Axe.csv",header=TRUE)
12 ListeCandidats=c("NAT","FR","EM","JL","MLP","EZ","JLM","AH","YJ","VP","PP","NDA")
13 base=as.data.frame(permutations(n=9,r=9,v=ListeCandidats[c(3,7,10,8,5,2,9,11,1)]))
14 base$V10=rep("EZ",120)
15 base$V11=rep("NDA",120)
16 base$V12=rep("JL",120)
17
18
19 base<-base[,c(which(colnames(base)=="V1"),which(colnames(base)=="V2"),which(colnames(base)=="V3"),which(colnames(base)=="V4"),which(

```

Code R définissant un axe avec 9 permutations et 3 candidats bloqués

Nous avons ensuite rangé les résultats que l'on obtenait pour 6 et 9 permutations dans une seule base, en nommant les axes avec 9 permutation par un H. Après avoir rangé les axes larges du plus petit au plus grand et les axes stricts du plus grand au plus petit, nous avons obtenu les résultats suivants.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	TTLarge	TTStrict	
2	105 NAT	PP	FR	AH	YJ	JLM	EM	VP	JL	MLP	EZ	NDA	2499	400	
3	110 NAT	PP	FR	AH	YJ	JLM	EM	VP	JL	MLP	EZ	NDA	2499	400	
4	261 NAT	PP	FR	AH	YJ	JLM	EM	VP	JL	MLP	EZ	NDA	2499	400	
5	H21 NAT	PP	FR	AH	YJ	JLM	EM	VP	JL	MLP	EZ	NDA	2499	400	
6	H81 NAT	PP	FR	AH	YJ	JLM	EM	VP	JL	MLP	EZ	NDA	2499	400	

Meilleur axe au sens large

Le meilleur axe au sens large compte donc 2499 trous. Nous avons obtenu cet axe avec le programme sur 6 permutations et le programme sur 9 permutations.

10170	H20 NAT	PP	JLM	FR	AH	YJ	EM	VP	MLP	EZ	NDA	JL	2960	486
10171	H70 NAT	PP	JLM	FR	AH	YJ	EM	VP	MLP	EZ	NDA	JL	2960	486

Meilleur axe au sens strict

Le meilleur axe au sens strict compte 486 bulletins sans trous. Nous avons obtenus cet axe uniquement grâce aux programmes sur 9 permutations.

3. Limites du mode de calculs

Les meilleurs axes ont été obtenus par approximation. En effet, nous n'avons pas pu calculer les 12 permutations car les calculs prenaient trop de temps. De fait, nous avons dû bloquer des candidats que l'on estimé proches les uns des autres. Nous n'avons donc pas la preuve par le calcul que les scores liés à ces axes soient les meilleurs possibles. De plus, nous avons obtenu des axes différents du meilleur axe alors que la différence du nombre de trous était infime.

6	H81 NAT	PP	FR	AH	YJ	JLM	EM	VP	JL	MLP	EZ	NDA	2499	400
7	261 NAT	PP	FR	AH	YJ	JLM	JL	EM	VP	MLP	EZ	NDA	2501	409

Le meilleur axe au sens large et le deuxième meilleur axe au sens large

De plus, le meilleur axe du sens strict place Jean Lassalle tout à droite de l'axe. En effet, avec cet indicateur, les « petits » candidats sont centrifugés car il est moins coûteux en terme de trous de décentrer un candidat peu approuvé qu'un candidat souvent approuvé. Hormis cela, les deux axes représentent bien les candidats habituellement classés à gauche, au centre et à droite.

C – Décomposition des inégalités en Nouvelle-Calédonie

En parallèle du travail sur les données d'expérimentations pour la présidentielle, j'ai aussi travaillé sur le un algorithme permettant d'obtenir la décomposition des

inégalités de niveau de vie en Nouvelle-Calédonie selon plusieurs critères, notamment la distance du foyer à la mine la plus proche.

1. Contexte

Cette étude a été réalisée dans le cadre d'une conférence pour la Journée Économique des Outre-Mer 2022 se déroulant le 11 et 12 octobre à Paris. Cette conférence a reçu le soutien du ministère des Outre-Mer ainsi que du CNRS. Nous avons travaillé en collaboration avec Frédéric Chantreuil, Professeur des universités à l'Université de Nouvelle-Calédonie. Le but de ce travail était d'obtenir une décomposition des inégalités de niveau de vie en fonction de plusieurs facteurs. Nous avons notamment étudié les différences de niveau de vie en fonction du genre, de la région, du type d'habitat ou encore de la distance entre le foyer et la mine la plus proche.

2. Présentation des données

Nous avons à notre disposition une base de données du recensement de la population Néo-Calédonienne, soit 204 721 individus. Grâce à celle-ci, nous avons accès à plusieurs variables quantitatives et qualitatives : Age, genre, commune,...

Pour notre étude, nous avons décidé de nous focaliser sur les variables suivantes :

- AGE : Age de l'individu
- S : Genre de l'individu
- ETM : Situation conjugale de l'individu (Célibataire ; Marié(e) ; Veuf.ve ; Divorcé(e))
- AGRI : Si l'individu est agriculteur (Non ; Oui consommation personnelle ; Oui vente)
- CSP : Catégorie socio-professionnelle de l'individu (Agriculteur ; Artisan ; Cadre ; Profession intermédiaire ; Employé ; Ouvrier ; Retraité, Autre inactif)
- DIP : Diplôme le plus élevé obtenu par l'individu (Aucun ; CEP ; BEPC ; CAP ; Bac général ; Bac techno ; Bac pro ; Université)

- DIST_MIN : Plus petite distance à la mine pour l'ensemble des individus selon CATIRIS
- CATIRIS : Numéro d'IRIS
- IDNV : Indice de niveau de vie de l'individu (voir annexe 2)

	A	B	C	D	E	F	G	H	I
1	CATIRIS	AGE	AGRI	CSP	DIP	ETM	S	IDNV	DIST_MIN
2	101	14	1 Non	8 Autre in	4 CAP	1 Celibata	2 Femme	16	28
3	101	14	1 Non	8 Autre in	1 Aucun	1 Celibata	2 Femme	8	28
4	101	14	1 Non	8 Autre in	1 Aucun	1 Celibata	2 Femme	15	28
5	101	14	1 Non	8 Autre in	1 Aucun	1 Celibata	2 Femme	13	28
6	101	14	2 Oui cons	8 Autre in	2 CEP	1 Celibata	1 Homme	16	28
7	101	14	1 Non	8 Autre in	2 CEP	1 Celibata	1 Homme	13	28
8	101	14	2 Oui cons	8 Autre in	2 CEP	1 Celibata	1 Homme	13	28
9	101	14	1 Non	8 Autre in	2 CEP	1 Celibata	1 Homme	14	28
10	101	14	1 Non	8 Autre in	2 CEP	1 Celibata	2 Femme	12	28
11	101	14	1 Non	8 Autre in	1 Aucun	1 Celibata	1 Homme	13	28
12	101	14	1 Non	8 Autre in	3 BEPC	1 Celibata	2 Femme	13	28
13	101	14	1 Non	8 Autre in	3 BEPC	1 Celibata	1 Homme	13	28
14	101	14	1 Non	8 Autre in	1 Aucun	1 Celibata	1 Homme	14	28
15	101	14	1 Non	8 Autre in	1 Aucun	1 Celibata	1 Homme	16	28
16	101	14	1 Non	8 Autre in	4 CAP	1 Celibata	1 Homme	13	28
17	101	14	1 Non	8 Autre in	4 CAP	1 Celibata	1 Homme	13	28
18	101	14	1 Non	8 Autre in	4 CAP	1 Celibata	2 Femme	14	28
19	101	14	1 Non	8 Autre in	4 CAP	1 Celibata	2 Femme	13	28
20	101	14	2 Oui cons	8 Autre in	1 Aucun	1 Celibata	2 Femme	12	28
21	101	15	1 Non	8 Autre in	1 Aucun	1 Celibata	1 Homme	14	28
22	101	15	1 Non	8 Autre in	1 Aucun	1 Celibata	1 Homme	15	28

Extrait de la base de données

3. Algorithme de décomposition des inégalités

Pour obtenir un pourcentage de contribution de chaque variable dans la composition des inégalités, nous avons utilisé l'Algorithme de Shapley. Nous avons repris l'algorithme écrit par Kévin Fourrey, Maître de Conférence à l'université Paris-Est Créteil. En résumé, l'algorithme forme des coalitions de une ou plusieurs variables et regarde si l'indice de niveau de vie moyen de la coalition formée est au-dessus ou en-dessous de toutes les autres coalitions formées. L'algorithme donne

en sortie le pourcentage de contribution de chaque variable à l'inégalité totale. Le pourcentage d'inégalité non expliquée par les variables ayant été choisies pour faire la décomposition est considéré comme un résidu. Les aspects techniques du calcul sont détaillés en annexe 3.

```

32 shapley_ineq <- function(distri, income,
33                           game = "equa",
34                           data_weights=F,
35                           index = c("Gini", "Entropy", "Atkinson", "Kolm"), theta =1,
36 ~                           get_csv=T, name_table="results") {
37
38   #Etape 0 : La désagrégation
39
40   #S'il n'y a pas de pondérations :
41 ~   if (data_weights==F){
42 ~     distri$weight = rep(1,nrow(distri)) }
43 ~   else {
44 ~     distri$weight = data_weights
45 ~   }
46
47   ds <- cbind(distri, income)
48
49   #La désagrégation se faire selon l'ordre des attributs dans la table ds
50
51   # average income per c1
52   ds <- ddply(ds, 1, mutate, Income_c1=weighted.mean(income, weight))
53
54   # average income per c1 and c2
55   ds <- ddply(ds, c(1,2), mutate, Income_c1_c2=weighted.mean(income, weight))
56
57   # average income per c1, c2 and c3
58   ds <- ddply(ds, c(1,2,3), mutate, Income_c1_c2_c3=weighted.mean(income, weight))
59
60   # average income per c1, c2, c3, and c4
61   ds <- ddply(ds, c(1,2,3,4), mutate, Income_c1_c2_c3_c4=weighted.mean(income, weight))
62
63   # average income per c1, c2, c3, c4 and c5
64   ds <- ddply(ds, c(1,2,3,4,5), mutate, Income_c1_c2_c3_c4_c5=weighted.mean(income, weight))
65
66   # average income per c1, c2, c3, c4, c5 and c6
67   ds <- ddply(ds, c(1,2,3,4,5,6), mutate, Income_c1_c2_c3_c4_c5_c6=weighted.mean(income, weight))
68
69   # average income per c1, c2, c3, c4, c5, c6 and c7
70   ds <- ddply(ds, c(1,2,3,4,5,6,7), mutate, Income_c1_c2_c3_c4_c5_c6_c7=weighted.mean(income, weight))
71
72   # average income per c1, c2, c3, c4, c5, c6 and c7
73   ds <- ddply(ds, c(1,2,3,4,5,6,7,8), mutate, Income_c1_c2_c3_c4_c5_c6_c7_c8=weighted.mean(income, weight))
74
75
76   # Differences between averages
77   ds$DIFF_c1_c2=ds$Income_c1_c2-ds$Income_c1
78   ds$DIFF_c1_c2_c3=ds$Income_c1_c2_c3-ds$Income_c1_c2
79   ds$DIFF_c1_c2_c3_c4=ds$Income_c1_c2_c3_c4-ds$Income_c1_c2_c3
80   ds$DIFF_c1_c2_c3_c4_c5=ds$Income_c1_c2_c3_c4_c5-ds$Income_c1_c2_c3_c4
81   ds$DIFF_c1_c2_c3_c4_c5_c6=ds$Income_c1_c2_c3_c4_c5_c6-ds$Income_c1_c2_c3_c4_c5
82   ds$DIFF_c1_c2_c3_c4_c5_c6_c7=ds$Income_c1_c2_c3_c4_c5_c6_c7-ds$Income_c1_c2_c3_c4_c5_c6
83   ds$DIFF_c1_c2_c3_c4_c5_c6_c7_c8=ds$Income_c1_c2_c3_c4_c5_c6_c7_c8-ds$Income_c1_c2_c3_c4_c5_c6_c7
84
85   # Value for residual :
86   ds$RESIDUAL=income-ds$Income_c1_c2_c3_c4_c5_c6_c7_c8

```

Début du code R de la fonction Shapley permettant la décomposition des inégalités

Source : Kévin Fourrey

A cause des limites techniques de l'ordinateur que nous avons à disposition, nous avons choisi de faire le calcul sur 7 variables, limitant à 127 le nombre de coalitions.

Il ne me restait plus qu'à modifier le programme R permettant de lancer la fonction.

```

1  install.packages(c('Rcpp','plyr','dplyr','GameTheoryAllocation',
2                      'GiniwegNeg'))
3
4  library(readxl)
5  library(Rcpp)
6  library(plyr)
7  library(dplyr)
8  library(GameTheoryAllocation)
9  library(GiniwegNeg)
10
11 data <- read_excel("C:/Data/MINES_NC/Base_NC.xlsx")
12 data <- as.data.frame(data)
13
14 #Il faut faire tourner le code de la fonction d'abord (cf ci-dessous)
15
16 data_select <- data %>%
17   select("AGE","S","ETM","AGRI","CSP","DIP","Dist_Min","CATIRIS")
18
19 decompo <- shapley_ineq(distri = data_select, income=data$IDNV,
20                         game = "equa",
21                         data_weights=F,
22                         index = "Gini",
23                         get_csv=T,name_table="results")
24
25 decompo$IMPORTANCE

```

Programme R permettant d'appliquer l'algorithme de Shapley à notre base de données

4. Résultats

Après plusieurs heures, nous obtenions en sortie un fichier csv avec la proportion d'inégalité expliquée pour chaque variable. Il ne restait plus qu'à multiplier par cent le résultat pour obtenir un pourcentage. Nous avons décidé d'obtenir les résultats selon la région des individus (Nord, Sud, Est, Ouest). En effet, la majorité de la population de la province Sud étant regroupée à Nouméa, nous voulions étudier séparément les effets de la distance à la mine sur l'inégalité en sachant que les habitants de Nouméa et donc de la province Sud habitent souvent loin des mines.

Après avoir séparé les résultats selon la province, nous obtenions le fichier ci-dessous.

	A	B	C	D	E	F	G	H	I	J	K
1	Base utilisé	AGE	S	ETM	AGRI	CSP	DIP	Dist_Min	CATIRIS	Residuals	Ineq
2	Nord	0.01893062	0.00794247	0.02915472	0.03920070	0.07566955	0.07304015	0.11456435	0.01639968	0.62509767	1
3	Sud	0.02432325	0.00419255	0.02043871	0.01137170	0.07990737	0.06769416	0.09891681	0.06503276	0.62812257	1
4	Est	0.01660944	0.00779860	0.02630430	0.02999315	0.07328141	0.07020811	0.11563381	0.01744361	0.64272750	1
5	Ouest	0.02480275	0.00418065	0.02021726	0.01149611	0.07927670	0.06801158	0.10043206	0.06677756	0.62480530	1

Environ 38 % des inégalités sont expliquées par les variables choisis. Comme nous l'avions prévu, le pourcentage d'inégalités expliqué par la distance à la mine est moindre dans la province Sud (9 %). Cela s'explique en partie par le fait que les habitants de la province Sud travaillent en majorité à Nouméa et sont donc moins concernés par le travail à la mine. Nous observons aussi que la catégorie socio-professionnelle et le niveau d'étude explique chacun environ 7 % des inégalités en Nouvelle-Calédonie et reste donc une préoccupation majeure pour les pouvoirs publics, tout comme la distance à la mine dans les provinces non-urbanisées, où cette variable explique 11 % des inégalités.

Conclusion

Pour conclure, ce stage au sein du centre de recherche en économie et management m'a permis de découvrir plusieurs aspects de la recherche en économie.

J'ai tout d'abord pu participer à des projets liés à l'actualité politique et sociale notamment à travers l'étude des systèmes de vote et la mesure des inégalités en Nouvelle-Calédonie. J'ai ainsi eu accès à des données réelles permettant d'effectuer un travail empirique.

De plus, j'ai eu l'opportunité de travailler en collaboration avec différents corps de métiers et de découvrir les bonnes pratiques pour travailler à distance comme organiser son répertoire de travail, faire état de l'avancée de mon travail régulièrement ou encore savoir communiquer mes résultats.

Ce stage m'aura aussi permis d'approfondir les concepts vus pendant mon année de Master, en développant des algorithmes statistiques sur R. J'ai notamment eu à utiliser les fonctions permettant de calculer une moyenne ou encore de construire un *data frame* à partir d'une base de données existante.

Enfin, ce stage a été l'occasion pour moi d'affiner mon projet professionnel et de découvrir les métiers liés à la recherche. J'ai pu apprécier la partie conception et communication de la recherche mais j'ai aussi découvert les aspects administratifs permettant au chercheur de s'impliquer dans la vie de son université et de son laboratoire.

Bibliographie

Sur les modes de scrutin :

Balinski et Laraki. *Majority judgment vs. Approval voting*. Operations Research 2020

Baujard et Lebon. *Le vote utile est-il un problème*, The conversation, 2022

Blais, Laslier, Sauger and Van der Straeten. *Sincere, strategic, and heuristic voting under four election rules : an experimental study with*. Social Choice and Welfare 35 : 435-472 (2009)

Chavalarias et Laslier. *Le jugement majoritaire peut-il sauver la démocratie ?* 27 janvier 2022, [Le drenché](#), Espace de Débat de Ouest France

Fabre. *Tie-breaking the highest median : alternatives to the majority judgment*. Social Choice and Welfare, 2021, 56 : 101–124

Igersheim, Baujard, Lebon. *Quel mode de scrutin pour quelle démocratie*. The conversation, 2022

Laslier. *And the looser is... Plurality voting*. In Electoral Systems edited by Dan Felsenthal and Moshé Machover, Springer, 2012

Laslier. *Lessons from In Situ experiments during French elections*. In *In Situ and Laboratory Experiments on Electoral Law Reform*, Dolez, Grofman, & Laurent (eds.), 2011, Springer

Laslier. *L'étrange « jugement majoritaire »* Revue économique, 2019, 70(4) : 569-588

Sur la décomposition des inégalités :

Chantreuil, F., Trannoy, A., 2011. *Inequality decomposition values*. Ann. Econ. Stat. 101–102, 6–29.

Chantreuil, F., Trannoy, A., 2013. *Inequality decomposition values: the trade-off between marginality and efficiency*. J. Econ. Inequal. 11 (1), 83–98.

Shapley, L.S., 1953. *A value for n-person game*. In: Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games. In: Annals of Mathematics Studies, vol. 2. Princeton University Press, Princeton, pp. 307–317

Annexes

Annexe 1 : Détails des questions posées aux personnes participant à l'expérimentation

Questionnaire de type A

ECRAN 1

VOTER AUTREMENT (A) TEST

Dans le contexte de l'élection présidentielle de 2022, une équipe de chercheurs teste différents modes de scrutin. Cette expérimentation prend environ 30 minutes. Sa finalité est purement scientifique. Il ne s'agit pas d'un sondage électoral.

Les données recueillies sont complètement anonymes.

Les modes de scrutin testés vous seront expliqués au fur et à mesure de l'expérience.

Cliquer sur SUIVANT pour commencer.

ECRAN 2

Etes-vous inscrit sur les listes électorales françaises ?

- 1 Oui
- 2 Non

ECRAN 3

QUI ETES-VOUS ?

Quel âge avez-vous ?

- 1 18-24 ans
- 2 25-34 ans
- 3 35-49 ans

- 4 50-64 ans
- 5 65 ans et plus

ECRAN 4

QUI ETES-VOUS ?

Quel est votre genre ?

- 6 Femme
- 7 Homme

ECRAN 5

QUI ETES-VOUS ?

Quel est votre catégorie socio-professionnelle ?

- 8 Agriculteur, artisan, profession libérale, chef d'entreprise
- 9 Cadre et profession intellectuelle supérieure
- 10 Profession intermédiaire
- 11 Employé
- 12 Ouvrier
- 13 Retraité
- 14 Inactif
- 15 Étudiant

ECRAN 6

QUI ETES-VOUS ?

De quelle région venez-vous ?

- 16 Auvergne-Rhône-Alpes
- 17 Bourgogne-Franche-Comté
- 18 Bretagne
- 19 Centre-Val de Loire
- 20 Corse
- 21 Grand Est
- 22 Hauts-de-France
- 23 Ile-de-France
- 24 Normandie
- 25 Nouvelle-Aquitaine
- 26 Occitanie
- 27 Pays de la Loire
- 28 Provence-Alpes-Côte d'Azur

ECRAN 7

QUI ETES-VOUS ?

Quelle est la taille de votre agglomération ?

- 29 Moins de 2 000 habitants
- 30 2 000 à 19 999 habitants
- 31 20 000 à 99 999 habitants
- 32 Plus de 100 000 habitants
- 33 Agglomération parisienne

QUESTION AVEC REPONSES DIFFERENTES

ECRAN 8

VOTEZ-VOUS ?

La méthode de vote officielle s'appelle le « scrutin uninominal majoritaire à deux tours ».

Chaque électeur vote pour un(e) candidat(e) unique. Gagne l'élection le ou la candidate qui obtient la majorité absolue au premier tour, ou l'emporte au second tour qui oppose les deux candidats ayant obtenu les scores les plus élevés au premier tour.

Allez-vous voter lors du premier tour de l'élection présidentielle française le 10 avril 2022 ?

- 34 Oui
- 35 Non
- 36 Je ne sais pas encore

Si OUI alors :

ECRAN 9

VOTEZ-VOUS ?

Si vous votiez aujourd'hui, quel serait votre vote au premier tour de l'élection présidentielle française ?

(Les candidats sont présentés dans l'ordre officiel du Conseil Constitutionnel)

- 37 Nathalie Arthaud
- 38 Fabien Roussel
- 39 Emmanuel Macron
- 40 Jean Lassalle
- 41 Marine Le Pen
- 42 Éric Zemmour
- 43 Jean-Luc Mélenchon
- 44 Anne Hidalgo
- 45 Yannick Jadot
- 46 Valérie Pécresse
- 47 Philippe Poutou
- 48 Nicolas Dupont-Aignan
- 49 Blanc
- 50 Nul

Si NON alors :

ECRAN 10

VOTEZ-VOUS ?

Vous allez vous abstenir (une seule réponse possible) :

- 51 Pour exprimer votre rejet de l'ensemble des candidats.
- 52 Parce que vous n'arrivez pas à choisir entre plusieurs candidats.
- 53 Parce que cette élection ne vous intéresse pas.
- 54 Parce que vous n'avez pas assez étudié les programmes des candidats pour faire un choix.
- 55 Parce que vous pensez que le résultat de cette élection ne va rien changer.
- 56 Autre : ...

Si JE NE SAIS PAS alors :

ECRAN 11

VOTEZ-VOUS ?

Au cas où vous iriez voter au premier tour de l'élection présidentielle française, quel serait votre vote ?

(Les candidats sont présentés dans l'ordre officiel du Conseil Constitutionnel)

- 57 Nathalie Arthaud
- 58 Fabien Roussel
- 59 Emmanuel Macron
- 60 Jean Lassalle
- 61 Marine Le Pen
- 62 Éric Zemmour
- 63 Jean-Luc Mélenchon
- 64 Anne Hidalgo
- 65 Yannick Jadot
- 66 Valérie Pécresse
- 67 Philippe Poutou
- 68 Nicolas Dupont-Aignan
- 69 Blanc
- 70 Nul

QUESTION AVEC REPONSES DIFFERENTES

ECRAN 12

VOTE PAR APPROBATION

Le vote par approbation est un mode de scrutin qui permet aux électeurs de se prononcer sur chacun des candidats en l'approuvant ou en ne l'approuvant pas.

L'électeur peut approuver (voter pour) autant de candidats qu'il le souhaite.

Avec le vote par approbation, le vainqueur de l'élection est celui ou celle qui obtient le plus grand nombre d'approbations.

71 J'ai bien compris ce mode de scrutin, je poursuis le questionnaire

72 Je voudrais plus d'explications sur ce mode de scrutin

Si PLUS EXPLICATIONS alors cet écran en plus :

ECRAN 13

VOTE PAR APPROBATION

Le score obtenu par un candidat correspond au nombre d'approbations qu'il a reçu.

Prenons l'exemple d'un cas dans lequel 2 électeurs E1 et E2 se prononcent sur 3 candidats C1, C2 et C3.

E1 vote de la façon suivante :

	Non approuvé	Approuvé
C1		X
C2	X	
C3		X

E2 vote de la façon suivante :

	Non approuvé	Approuvé
C1	X	
C2		X
C3		X

Les résultats sont donc les suivants :

Candidats	Nombre d'approbations
C1	1

C2	1
C3	2

C'est donc C3 qui l'emporte avec 2 approbations.

ECRAN 14

VOTE PAR APPROBATION

A votre avis, avec ce mode de scrutin, comment un électeur qui veut apporter son soutien à un candidat X, va-t-il voter ?

	Non approuvé	Approuvé
Candidat X	<input type="radio"/>	<input type="radio"/>

ECRAN 15

VOTE PAR APPROBATION

Quels sont les candidats(e)s de l'élection présidentielle de 2022 que vous approuveriez ?

(Les candidats sont présentés dans l'ordre officiel du Conseil Constitutionnel)

Indication : Pour le résultat du vote, lorsqu'un électeur ne se prononce pas sur un candidat, on considère qu'il ne l'approuve pas.

	Non approuvé	Approuvé
Nathalie Arthaud	<input type="radio"/>	<input type="radio"/>
Fabien Roussel	<input type="radio"/>	<input type="radio"/>
Emmanuel Macron	<input type="radio"/>	<input type="radio"/>
Jean Lassalle	<input type="radio"/>	<input type="radio"/>
Marine Le Pen	<input type="radio"/>	<input type="radio"/>
Éric Zemmour	<input type="radio"/>	<input type="radio"/>
	Non approuvé	Approuvé
Jean-Luc Mélenchon	<input type="radio"/>	<input type="radio"/>
Anne Hidalgo	<input type="radio"/>	<input type="radio"/>

Yannick Jadot
Valérie Pécresse
Philippe Poutou
Nicolas Dupont-Aignan

°
°
°
°

°
°
°
°

QUESTION AVEC REPONSES DIFFERENTES :

ECRAN 16

VOTE PAR NOTE

Le vote par note est un mode de scrutin qui permet aux électeurs de s'exprimer sur chacun des candidats en lui attribuant une note sur l'échelle proposée (-1 ;0 ;1 ;2 ;3 ;4).

Les notes attribuées à chaque candidat sont additionnées.

Le vainqueur est celui ou celle dont la somme des notes est la plus élevée, c'est-à-dire celui ou celle qui a la plus forte note moyenne.

73 J'ai bien compris ce mode de scrutin, je poursuis le questionnaire

74 Je voudrais plus d'explications sur ce mode de scrutin

Si PLUS EXPLICATIONS alors cet écran en plus :

ECRAN 17

VOTE PAR NOTE

Le score obtenu par un candidat correspond à la somme des notes qu'il a reçues.

Prenons l'exemple d'un cas dans lequel 2 électeurs E1 et E2 se prononcent sur 3 candidats C1,C2 et C3.

E1 vote de la façon suivante :

	-1	0	1	2	3	4
C1		X				

C2				X		
C3						X

E2 vote de la façon suivante :

	-1	0	1	2	3	4
C1						X
C2		X				
C3				X		

Les résultats sont donc les suivants :

Candidats	Score total	Note moyenne
C1	0+4=4	2
C2	2+0=2	1
C3	4+2=6	3

Note : Le point noir erreur pendant la capture d'écran

ECRAN 18

VOTE PAR NOTE

A votre avis, avec ce mode de scrutin, comment un électeur qui veut apporter un soutien maximal à un candidat X, va-t-il voter ?

	-1	0	1	2	3	4
Candidat X	°	°	°	°	°	°

ECRAN 19

VOTE PAR NOTE

Quelle note donneriez-vous à chacun(e) des candidat(e)s à l'élection présidentielle de 2022 si le mode de scrutin était le vote par note ?

Indication : En cas de non réponse sur un candidat, le note -1 lui est attribué par défaut.

	-1	0	1	2	3	4
Nathalie Arthaud	°	°	°	°	°	°

Fabien Roussel	°	°	°	°	°	°
Emmanuel Macron	°	°	°	°	°	°
Jean Lassalle	°	°	°	°	°	°
Marine Le Pen	°	°	°	°	°	°
Éric Zemmour	°	°	°	°	°	°
	-1	0	1	2	3	4
Jean-Luc Mélenchon	°	°	°	°	°	°
Anne Hidalgo	°	°	°	°	°	°
Yannick Jadot	°	°	°	°	°	°
Valérie Pécresse	°	°	°	°	°	°
Philippe Poutou	°	°	°	°	°	°
Nicolas Dupont-Aignan	°	°	°	°	°	°

ECRAN 20

SATISFACTION

Vous venez de faire l'expérience de trois modes de scrutin différents pour élire le président de la République. Quelle règle avez-vous préféré sur les trois ?

- 75 Le scrutin uninominal à deux tours (la règle officielle)
- 76 Le vote par approbation
- 77 Le vote par note

ECRAN 21

OPINION

Ceci n'est pas une procédure de vote, juste une façon de recueillir votre opinion sur les candidats de l'élection présidentielle de 2022.

Évaluez chacun de ces 12 candidats sur leur capacité à faire un bon président de la république en leur donnant un score entre 0 à 20 en cochant la case correspondante.

0 est le plus mauvais score, 20 est le meilleur. Si vous ne pouvez pas ou ne souhaitez pas vous exprimer sur un(e) candidat(e), cochez la case « PAS » signifiant « Je n'ai pas d'opinion ».

Indication : si vous n'indiquez pas d'opinion sur un(e) candidat(e), on suppose que vous êtes « Sans opinion » sur elle ou lui.

	PA	0	1	2	3	4	5	6	7	8	9	1	1	1	1	1	1	1	1	1	2	
	S											0	1	2	3	4	5	6	7	8	9	0
Nathalie	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°

Arthaud																						
Fabien	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Roussel																						
Emmanue	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
I Macron																						
Jean	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Lassalle																						
Marine Le	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Pen																						
Éric	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Zemmour																						
	PA	0	1	2	3	4	5	6	7	8	9	1	1	1	1	1	1	1	1	1	1	2
	S											0	1	2	3	4	5	6	7	8	9	0
Jean-Luc	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Mélencho																						
n																						
Anne	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Hidalgo																						
Yannick	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Jadot																						
Valérie	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Pécresse																						
Philippe	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Poutou																						
Nicolas	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°	°
Dupont-																						
Aignan																						

ECRAN 22

APPROFONDISSEMENT

On considère une élection à laquelle se présentent 2 candidats : le candidat A et le candidat B

Chacun des 11 électeurs s'est prononcé sur 2 candidats en lui donnant une note. Il en résulte la répartition ci-dessous des notes attribuées chacun des deux candidats (chaque croix représente un vote) :

	-1	0	1	2	3	4
Candidat A	XX	XX	X	XX	XXX	X
Candidat B	X	X	XXXX	X	X	XXX

Avec le vote par note, quel est selon vous le candidat qui a été élu par ces 11 électeurs ?

78 Candidat A

79 Candidat B

ECRAN 23

APPROFONDISSEMENT

11 électeurs se prononcent par approbation sur les quatres candidats respectivement désignés par A, B, C et D. (chaque croix représente un vote)

	Non approuvé	Approuvé
Candidat A	XXXXXX	XXXXX
Candidat B	XXXX	XXXXXXXX
Candidat C	XXXXXXXXX	XXX
Candidat D	XXXXXXXX	XXXX

Avec le vote par approbation, quel est selon vous le candidat qui a été élu par ces 11 électeurs ?

- 80 Candidat A
- 81 Candidat B
- 82 Candidat C
- 83 Candidat D

ECRAN 24

APPROFONDISSEMENT

Sachant que le(la) candidat(e) qui a la meilleure note est élu(e), le vote par note attribue à chaque candidat(e) une note qui :

	Vrai	Faux	Je ne sais pas
Correspond à la note moyenne donnée à ce candidat par l'ensemble des électeurs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correspond à la note moyenne donnée à ce candidat par la moitié plus un des électeurs qui lui sont le plus favorables	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correspond à la note donnée par l'électeur qui partage en deux l'ensemble des électeurs classés du plus favorable au plus défavorable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

ECRAN 25

APPROFONDISSEMENT

Quelle est votre position sur l'affirmation suivante : « les citoyens doivent pouvoir comprendre/déduire facilement le résultat de l'élection après le dépouillement, une fois connu le contenu des bulletins ». C'est :

- 84 Inutile
- 85 Souhaitable
- 86 Important
- 87 Fondamentale/Indispensable
- 88 Autre : ...

Questionnaire de type B

ECRAN 1

VOTER AUTREMENT (B) TEST

Dans le contexte de l'élection présidentielle de 2022, une équipe de chercheurs teste différents modes de scrutin. Cette expérimentation prend environ 30 minutes. Sa finalité est purement scientifique. Il ne s'agit pas d'un sondage électoral.

Les données recueillies sont complètement anonymes.

Les modes de scrutin testés vous seront expliqués au fur et à mesure de l'expérience.

Cliquer sur SUIVANT pour commencer.

ECRAN 2

Etes-vous inscrit sur les listes électorales françaises ?

- 3 Oui
- 4 Non

ECRAN 3

QUI ETES-VOUS ?

Quel âge avez-vous ?

- 89 18-24 ans
- 90 25-34 ans
- 91 35-49 ans
- 92 50-64 ans
- 93 65 ans et plus

ECRAN 4

QUI ETES-VOUS ?

Quel est votre genre ?

- 94 Femme
- 95 Homme

ECRAN 5

QUI ETES-VOUS ?

Quel est votre catégorie socio-professionnelle ?

- 96 Agriculteur, artisan, profession libérale, chef d'entreprise
- 97 Cadre et profession intellectuelle supérieure
- 98 Profession intermédiaire
- 99 Employé
- 100 Ouvrier
- 101 Retraité
- 102 Inactif
- 103 Étudiant

ECRAN 6

QUI ETES-VOUS ?

De quelle région venez-vous ?

- 104 Auvergne-Rhône-Alpes

- 105 Bourgogne-Franche-Comté
- 106 Bretagne
- 107 Centre-Val de Loire
- 108 Corse
- 109 Grand Est
- 110 Hauts-de-France
- 111 Ile-de-France
- 112 Normandie
- 113 Nouvelle-Aquitaine
- 114 Occitanie
- 115 Pays de la Loire
- 116 Provence-Alpes-Côte d’Azur

ECRAN 7

QUI ETES-VOUS ?

Quelle est la taille de votre agglomération ?

- 117 Moins de 2 000 habitants
- 118 2 000 à 19 999 habitants
- 119 20 000 à 99 999 habitants
- 120 Plus de 100 000 habitants
- Agglomération parisienne

QUESTION AVEC REPONSES DIFFERENTES

ECRAN 8

VOTEZ-VOUS ?

La méthode de vote officielle s'appelle le « scrutin uninominal majoritaire à deux tours ».

Chaque électeur vote pour un(e) candidat(e) unique. Gagne l'élection le ou la candidate qui obtient la majorité absolue au premier tour, ou l'emporte au second tour qui oppose les deux candidats ayant obtenu les scores les plus élevés au premier tour.

Allez-vous voter lors du premier tour de l'élection présidentielle française le 10 avril 2022 ?

121 Oui

122 Non

123 Je ne sais pas encore

Si OUI alors :

ECRAN 9

VOTEZ-VOUS ?

Si vous votiez aujourd'hui, quel serait votre vote au premier tour de l'élection présidentielle française ?

(Les candidats sont présentés dans l'ordre officiel du Conseil Constitutionnel)

124 Nathalie Arthaud

125 Fabien Roussel

126 Emmanuel Macron

127 Jean Lassalle

128 Marine Le Pen

129 Éric Zemmour

130 Jean-Luc Mélenchon

131 Anne Hidalgo

132 Yannick Jadot

133 Valérie Pécresse

134 Philippe Poutou

135 Nicolas Dupont-Aignan

136 Blanc

137 Nul

Si NON alors :

ECRAN 10

VOTEZ-VOUS ?

Vous allez vous abstenir (une seule réponse possible) :

- 138 Pour exprimer votre rejet de l'ensemble des candidats.
- 139 Parce que vous n'arrivez pas à choisir entre plusieurs candidats.
- 140 Parce que cette élection ne vous intéresse pas.
- 141 Parce que vous n'avez pas assez étudié les programmes des candidats pour faire un choix.
- 142 Parce que vous pensez que le résultat de cette élection ne va rien changer.
- 143 Autre : ...

Si JE NE SAIS PAS alors :

ECRAN 11

VOTEZ-VOUS ?

Au cas où vous iriez voter au premier tour de l'élection présidentielle française, quel serait votre vote ?

(Les candidats sont présentés dans l'ordre officiel du Conseil Constitutionnel)

- 144 Nathalie Arthaud
- 145 Fabien Roussel
- 146 Emmanuel Macron
- 147 Jean Lassalle
- 148 Marine Le Pen
- 149 Éric Zemmour
- 150 Jean-Luc Mélenchon
- 151 Anne Hidalgo
- 152 Yannick Jadot
- 153 Valérie Pécresse
- 154 Philippe Poutou
- 155 Nicolas Dupont-Aignan
- 156 Blanc

QUESTION AVEC REPONSES DIFFERENTES

ECRAN 12

VOTE PAR APPROBATION

Le vote par approbation est un mode de scrutin qui permet aux électeurs de se prononcer sur chacun des candidats en l'approuvant ou en ne l'approuvant pas.

L'électeur peut approuver (voter pour) autant de candidats qu'il le souhaite.

Avec le vote par approbation, le vainqueur de l'élection est celui ou celle qui obtient le plus grand nombre d'approbations.

158 J'ai bien compris ce mode de scrutin, je poursuis le questionnaire

159 Je voudrais plus d'explications sur ce mode de scrutin

Si PLUS EXPLICATIONS alors cet écran en plus :

ECRAN 13

VOTE PAR APPROBATION

Le score obtenu par un candidat correspond au nombre d'approbations qu'il a reçu.

Prenons l'exemple d'un cas dans lequel 2 électeurs E1 et E2 se prononcent sur 3 candidats C1, C2 et C3.

E1 vote de la façon suivante :

	Non approuvé	Approuvé
C1		X
C2	X	
C3		X

E2 vote de la façon suivante :

	Non approuvé	Approuvé
C1	X	
C2		X
C3		X

Les résultats sont donc les suivants :

Candidats	Nombre d'approbations
C1	1
C2	1
C3	2

C'est donc C3 qui l'emporte avec 2 approbations.

ECRAN 14

VOTE PAR APPROBATION

A votre avis, avec ce mode de scrutin, comment un électeur qui veut apporter son soutien à un candidat X, va-t-il voter ?

	Non approuvé	Approuvé
Candidat X	<input type="radio"/>	<input type="radio"/>

ECRAN 15

VOTE PAR APPROBATION

Quels sont les candidats(e)s de l'élection présidentielle de 2022 que vous approuveriez ?

(Les candidats sont présentés dans l'ordre officiel du Conseil Constitutionnel)

Indication : Pour le résultat du vote, lorsqu'un électeur ne se prononce pas sur un candidat, on considère qu'il ne l'approuve pas.

	Non approuvé	Approuvé
Nathalie Arthaud	<input type="radio"/>	<input type="radio"/>
Fabien Roussel	<input type="radio"/>	<input type="radio"/>
Emmanuel Macron	<input type="radio"/>	<input type="radio"/>
Jean Lassalle	<input type="radio"/>	<input type="radio"/>
Marine Le Pen	<input type="radio"/>	<input type="radio"/>
Éric Zemmour	<input type="radio"/>	<input type="radio"/>
	Non approuvé	Approuvé

Jean-Luc Mélenchon	°	°
Anne Hidalgo	°	°
Yannick Jadot	°	°
Valérie Pécresse	°	°
Philippe Poutou	°	°
Nicolas Dupont-Aignan	°	°

QUESTION AVEC REPONSES DIFFERENTES :

ECRAN 16

LE JUGEMENT MAJORITAIRE

L'électeur évalue chacun des candidats en lui attribuant l'une des mentions proposées : A rejeter ; Insuffisant ; Passable ; Assez Bien ; Bien ; Très bien.

A l'issue du vote, on calcule la « mention majoritaire » de chaque candidat, c'est-à-dire la mention médiane, telle qu'au moins 50% des votants lui ont attribué cette mention ou une meilleure mention. Le candidat élu est celui qui obtient la meilleure « mention majoritaire ».

Lorsque plusieurs candidats ont la même meilleure mention majoritaire, on départage les candidats en considérant que le plus

160 J'ai bien compris ce mode de scrutin, je poursuis le questionnaire

161 Je voudrais plus d'explications sur ce mode de scrutin

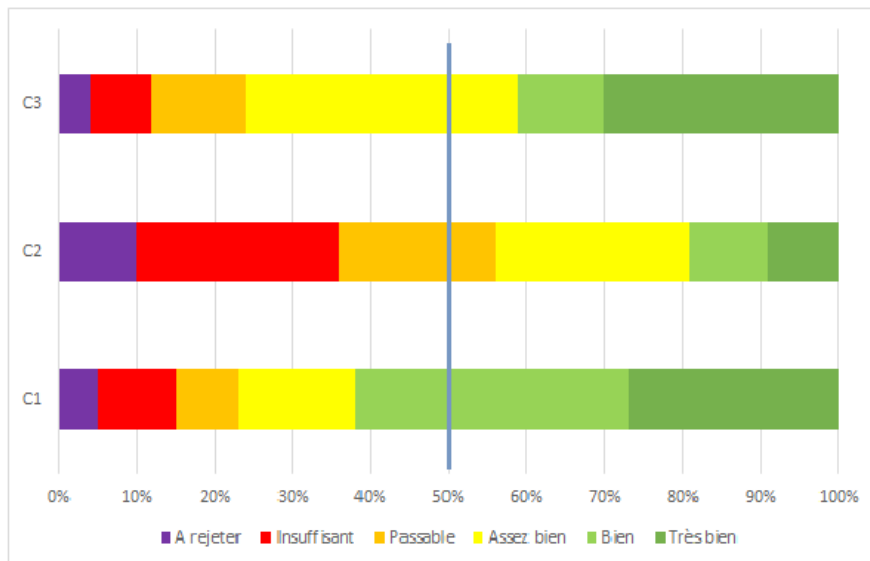
Si PLUS EXPLICATIONS alors cet écran en plus :

ECRAN 17

LE JUGEMENT MAJORITAIRE

Le résultat obtenu par un candidat correspond à la médiane des mentions qu'il a reçues. Le candidat qui a la meilleure l'emporte.

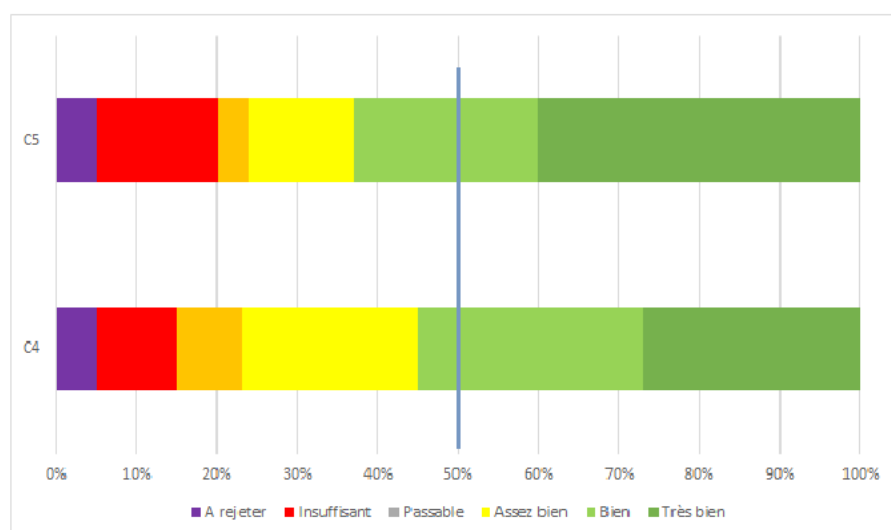
Prenons l'exemple d'un cas dans lequel les électeurs qui se prononcent sur 3 candidats C1, C2 et C3. La répartition des mentions qui leur sont attribuées, est représentée par les profils de mérite suivants :



Cette situation conduit au résultat suivant : le candidat C1 gagne l'élection car il obtient une meilleure mention majoritaire (Bien), soit mieux que le candidat C3 (Assez Bien) et le candidat C2 (Passable).

Que se passe-t-il si deux candidats obtiennent la même mention majoritaire ?

Prenons l'exemple d'un cas dans lequel les électeurs se sont prononcés sur 2 candidats C4 et C5 et leur ont donné la même mention majoritaire avec les profils de mérite suivants :



Dans cet exemple, C4 et C5 ont la mention majoritaire « Bien » tous les 2. Pour C4, les électeurs en désaccord avec la mention majoritaire sont plus nombreux à lui préférer une moins bonne mention qu'une meilleure mention. Pour C5, les électeurs en désaccord avec la mention majoritaire sont plus nombreux à lui préférer une meilleure mention qu'une moins bonne mention.

C'est donc le candidat C5 qui l'emporte.

ECRAN 18

LE JUGEMENT MAJORITAIRE

A votre avis, avec ce mode de scrutin, comment un électeur qui veut apporter un soutien maximal à un candidat X, va-t-il voter ?

Candidat X	A rejeter °	Insuffisant °	Passable °	Assez bien °	Bien °	Très bien °
------------	----------------	------------------	---------------	-----------------	-----------	----------------

ECRAN 19

LE JUGEMENT MAJORITAIRE

Quelle mention donneriez-vous à chacun(e) des candidat(e)s à l'élection présidentielle de 2022 si le mode de scrutin était le jugement majoritaire ?

Indication : En cas de non réponse sur un candidat, la mention « A rejeter » lui est attribué par défaut.

	A rejeter	Insuffisant	Passable	Assez bien	Bien	Très bien
Nathalie Arthaud	°	°	°	°	°	°
Fabien Roussel	°	°	°	°	°	°
Emmanuel Macron	°	°	°	°	°	°
Jean Lassalle	°	°	°	°	°	°
Marine Le Pen	°	°	°	°	°	°
Éric Zemmour	°	°	°	°	°	°
	A rejeter	Insuffisant	Passable	Assez bien	Bien	Très bien
Jean-Luc Mélenchon	°	°	°	°	°	°
Anne Hidalgo	°	°	°	°	°	°
Yannick Jadot	°	°	°	°	°	°
Valérie Pécresse	°	°	°	°	°	°
Philippe Poutou	°	°	°	°	°	°
Nicolas Dupont-Aignan	°	°	°	°	°	°

ECRAN 20

SATISFACTION

Vous venez de faire l'expérience de trois modes de scrutin différents pour élire le président de la République. Quelle règle avez-vous préféré sur les trois ?

162 Le scrutin uninominal à deux tours (la règle officielle)

163 Le vote par approbation

164 Le jugement majoritaire

ECRAN 21

OPINION

Ceci n'est pas une procédure de vote, juste une façon de recueillir votre opinion sur les candidats de l'élection présidentielle de 2022.

Évaluez chacun de ces 12 candidats sur leur capacité à faire un bon président de la république en leur donnant un score entre 0 à 20 en cochant la case correspondante.

0 est le plus mauvais score, 20 est le meilleur. Si vous ne pouvez pas ou ne souhaitez pas vous exprimer sur un(e) candidat(e), cochez la case « PAS » signifiant « Je n'ai pas d'opinion ».

Indication : si vous n'indiquez pas d'opinion sur un(e) candidat(e), on suppose que vous êtes « Sans opinion » sur elle ou lui.

	PA	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Nathalie Arthaud	S																					
Fabien Roussel																						
Emmanuelle Macron																						
Jean Lassalle																						
Marine Le Pen																						
Éric Zemmour																						
	PA	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	S																					

	Non approuvé	Approuvé
Candidat A	XXXXXX	XXXXX
Candidat B	XXXX	XXXXXXXX
Candidat C	XXXXXXXX	XXX
Candidat D	XXXXXX	XXXX

Avec le vote par approbation, quel est selon vous le candidat qui a été élu par ces 11 électeurs ?

- 167 Candidat A
- 168 Candidat B
- 169 Candidat C
- 170 Candidat D

ECRAN 24

APPROFONDISSEMENT

Sachant que le(la) candidat(e) qui a la meilleure mention est élu(e), le jugement majoritaire attribue à chaque candidat(e) une mention qui :

	Vrai	Faux	Je ne sais pas
Correspond à la mention moyenne donnée à ce candidat par l'ensemble des électeurs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correspond à la mention moyenne donnée à ce candidat par la moitié plus un des électeurs qui lui sont le plus favorables	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correspond à la mention donnée par l'électeur qui partage en deux l'ensemble des électeurs classés du plus favorable au plus défavorable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

ECRAN 25

APPROFONDISSEMENT

Quelle est votre position sur l'affirmation suivante : « les citoyens doivent pouvoir comprendre/déduire facilement le résultat de l'élection après le dépouillement, une fois connu le contenu des bulletins ». C'est :

- 171 Inutile
- 172 Souhaitable
- 173 Important
- 174 Fondamentale/Indispensable
- 175 Autre : ...

Annexe 2 : Détail de la construction de l'indicateur de niveau de vie

1.2. L'INDICATEUR DE NIVEAU DE VIE

Pour étudier l'effet localisé des mines sur le niveau de vie moyen des ménages il faut tout d'abord créer la variable dépendante : l'indicateur de niveau de vie.

N'ayant pas de données sur les dépenses des ménages ou leurs revenus, nous allons construire un indicateur de niveau de vie à partir des caractéristiques de logement et les équipements du ménage. Pour cela, nous utilisons les données issues du recensement 2014 évoquées plus haut.²

¹ Sans compter les communes qui composent le grand Nouméa, et donc plus peuplées que le reste de l'île aussi.

² Données transmises par l'ISEE dans le cadre du projet de recherche CNRT : Mines et Territoires.

Nous disposons de deux bases de données ;

Une base de données structurée par logement, avec ses caractéristiques et les biens d'équipement dont dispose le ménage, et

Une base individuelle avec les réponses relatives aux niveaux d'éducation des individus, leur âge, situation maritale, situation d'emploi, pratique d'activités annexes...

Dans un premier temps, je crée un indicateur qui prend en compte les variables de la colonne (1) (que je recode pour qu'elles soient ordonnées selon l'augmentation de niveau de vie qu'elles traduisent). Les variables utilisées sont similaires aux variables que Kobiane (2004) utilise pour mesurer le niveau de vie à Ouagadougou.

La création de l'indicateur demande une réflexion sur chaque modalité afin que l'indicateur final retranscrive le mieux possible le niveau de vie. Par exemple, si on s'intéresse au type de construction, il y a 5 modalités : maison, appartement, case, cabane, autres. Il serait tentant de se dire qu'avoir une maison rend compte d'un niveau de vie plus important qu'un appartement ou une case. Cependant, dans la réalité cela rend compte plutôt d'autres facteurs comme le lieu de vie (Nouméa ou en Brousse), le mode de vie et la communauté d'appartenance plutôt que la qualité de vie que l'on veut mesurer. En revanche on peut plus facilement opposer une de ces trois modalités à vivre en « cabane » (squat, bidonville).

Ainsi, comme mis en évidence dans la colonne « Modalités après recodage » du tableau 1, à chaque variable est attribuée un score entre 0 et 4. Faire simplement une somme donne la même importance à chaque caractéristique, disposer d'un réfrigérateur ou d'un téléphone mobile ne retranscrivent pas les mêmes niveaux de vie. Pour avoir un meilleur indicateur, rendant compte des différentes importances des variables selon la structure des données, nous pouvons utiliser l'analyse en composante principale (ACP)³ après avoir normalisé les différentes modalités. On rapporte ensuite l'indicateur sur un intervalle de 0 à 100, ce qui permettra une interprétation plus lisible.

Le tableau 1 décrit le poids de chaque variable sur l'indicateur retenu. Cet indicateur (comme toute variable utilisée comme proxy) n'est pas parfait et comporte un certain nombre de biais et de limite. Cet aspect sera discuté dans la section 5.

Tableau 1 :

<i>Nom</i>	<i>Intitulé</i>	<i>Modalités</i>	<i>Modalités après recodage</i>	<i>Poids des variables</i>
<i>BATO</i>	Nombre de bateau dans le ménage	1, 2, 3, 4, 5, 6, 7, 8 ou plus	0/1/2/3	0,0594
<i>MOTO</i>	Nombre de moto dans le ménage	1, 2, 3, 4, 5, 6, 7, 8 ou plus	0/1/2/3	0,0767
<i>TMOB</i>	Téléphone mobile	Oui / Non	0/1	0,137
<i>CHAUF</i>	Chauffage	Oui / Non	0/1	0,1576

³ " L'ACP est une méthode d'analyse factorielle dont l'objectif est la réduction d'un ensemble de variables quantitatives liées entre elles en un nombre restreint de facteurs ou composantes principales. La première composante dans une ACP explique presque toujours une très forte proportion de la variance et par conséquent peut être interprétée comme l'indicateur le plus synthétique des variables de base, donc dans notre cas le proxy de niveau de vie. " Kobiane ()

<i>SURF</i>	Surface (classes)	0,1, 2, 3, 4, 5,	0/1/2/3/4	0,189
<i>EAU</i>	Alimentation principale en eau	Eau courante / Point d'eau extérieur (individuel) / Point d'eau collectif	0/1/2	0,2017
<i>TC</i>	Type de construction	Maison / Appartement / Case / Cabane /Autre	0/2	0,204
<i>IS</i>	Baignoire ou douche	Oui / Non	0/1	0,2508
<i>WC</i>	WC	Oui / Non	0/1	0,2843
<i>VOIT</i>	Nombre de voiture dans le ménage	1, 2, 3, 4, 5, 6, 7, 8 ou plus	0/1/2/3	0,2866
<i>ELEC</i>	Mode principal d'éclairage	Réseau général d'électricité / Groupe électrogène / Panneaux solaires / Autre	0/1/2	0,2867
<i>CLIM</i>	Clim	Oui / Non	0/1	0,2885
<i>REFRI</i>	Frigo	Oui / Non	0/1	0,3109
<i>ORDI</i>	Ordinateur	Oui / Non	0/1	0,3247
<i>TFIXE</i>	Téléphone fixe	Oui / Non	0/1	0,3258
<i>INTER</i>	Internet	Oui / Non	0/1	0,3598

Annexe 3 : Éléments théorique de calculs d'une valeur de Shapley

(Source : Lebon Isabelle, Chantreuil Frédéric, *Gender contribution to income inequality*, Economics Letters 133 (2015) 27-30)

2. Gender decomposition

To decompose wage inequality by gender we consider the diverse dimensions of the status of an individual in terms of wage sources. In other words, the characteristics of an individual (age, occupation, gender, etc.) are treated as different sources for the wage. Consequently, for a given wage distribution and a set of individual's characteristics, the decomposition of the overall wage inequality can be effected using the Shapley approach. The contribution of a characteristic j to the overall wage inequality can be defined by the following formula:

$$Sh_j = \sum_{S \subset K, j \in S} \frac{(s-1)!(k-s)!}{k!} [I(Y(S)) - I(Y(S - \{j\}))] \quad (1)$$

where I is the chosen inequality measure (Gini and Theil in our following illustration), K is the set of individual's characteristics, k the cardinality of K , S a subset of individual's characteristics, s the cardinality of S , and $Y(S)$ the distribution of wage among wage sources obtained, equalizing the wage from the complementary source $N - S$ for all individuals. This distribution is defined by $Y(\emptyset) = 0$, and for all $S \in 2^K$, $S \neq \emptyset$,

$$Y(S) = \left(\sum y_1^j + \sum \mu(y^j), \dots, \sum y_n^j + \sum \mu(y^j) \right) \quad (2)$$

The total wage of this individual comprises the average wage of the agents of the same age of the subpopulation (of size n_a) considered:

$$w_a = \sum_{i=1}^{n_a} \frac{w_{i,a}}{n_a} \quad (4)$$

the difference between this average wage and those of the women/men of the same age (in number $n_{a,g}$) of the subpopulation considered:

$$w_{a,g} - w_a = \sum_{i=1}^{n_{a,g}} \frac{w_{i,a,g}}{n_{a,g}} - w_a \quad (5)$$

and a part specific to this woman/man i :

$$r_{i,a,g} = w_{i,a,g} - w_{a,g} \quad (6)$$

where $a \in [\underline{a}, \bar{a}]$ and $g = f, m$.

The first element of the wage described by Eq. (3) takes a strictly positive value. The second element is positive or negative according to whether the women/men of age a earned on average more or less than the agents of the same age in the subpopulation considered. The third element can be positive or negative according to whether agent i gains more or less than the average of the women/men of her/his age pertaining to the same subpopulation.

The wage sources being defined, we apply the Shapley decomposition method presented previously (Eq. (1)), which allows us to determine the contribution of age (considered as a proxy of career stage); gender; and the third variable which represents all other characteristics arising from the total wage inequality for a given population.

Other observable characteristics, like the diploma or the activity sector, can be used to group individuals into more homogeneous sub-populations. The decomposition of inequality into these more homogeneous sub-populations allows to better specify the part attributable to gender by omitting the structural effects which can arise if men and women have different characteristics.