

MUNICIPALES 2020 SUR TWITTER

Rapport national

Les transports et l'écologie dominant les échanges.
Paris imprime sa prédominance.
Les partis politiques sont peu mentionnés.

27 février 2020

KANTAR



**LA
VOX
POPULI**

Le Twittorama
des municipales 2020

SOMMAIRE

- 3 Préambule
- 4 Chiffres clés
- 6 Principaux enseignements

Focus sur les thématiques de la campagne

- 8 Les mots de la première partie de la campagne
- 9 Les thèmes de la première partie de la campagne
- 11 Classement des thématiques d'importance prioritaire pour les Français
- 13 Les thèmes les plus cités dans les 11 villes étudiées
- 14 Résultats des thèmes par ville
- 15 Des élections intermédiaires pour un scrutin nationalisé ?
La référence à Emmanuel Macron dans les échanges

Focus sur le podium des têtes de liste

- 18 Les têtes de liste les plus citées dans les 11 villes
- 19 Le top 30 des candidats les plus cités dans leurs villes
- 20 Les citations des partis politiques

Méthodologie



PRÉAMBULE

Espace d'expression caractérisé par sa spontanéité et sa rapidité de diffusion, Twitter apparaît comme un véhicule puissant des opinions, revendications ou contestations des citoyens sur la toile. On ne compte plus les exemples où quelques hashtags ont changé le cours des événements, voir celui de l'Histoire. Les Printemps arabes, le mouvement #MeToo ou plus récemment encore, l'affaire Griveaux l'ont amplement démontré. Ces phénomènes médiatiques ont mis en lumière l'influence croissante de Twitter, notamment dans les débats démocratiques.

Il nous est donc apparu nécessaire, à l'occasion de la campagne des Municipales 2020, de regarder de près ce qui allait se passer sur Twitter pour savoir si le réseau social était un reflet fidèle des préoccupations des Français ou s'il s'agissait plutôt d'un miroir déformant (mais néanmoins agissant). **C'est pourquoi La Vox Populi a mené, en partenariat avec Kantar, une analyse sémantique des tweets émis dans le cadre cette campagne, à Paris et dans les 10 plus grandes villes de France.**

L'étude vise à faire ressortir les sujets et objets du débat et de montrer la manière dont ils sont traités sur Twitter : quelles thématiques sont abordées dans chaque ville et nationalement ? Quelles têtes de liste sont associées à chaque thématique ? De qui parle-t-on ? Pour dire quoi et de quelle manière ?

Les enseignements de cette étude font l'objet d'un rapport complet sur chaque ville traitée, ainsi que d'une analyse nationale permettant de mettre en relief les particularités de la campagne dans chaque ville, avec quelques surprises montrant les spécificités de Twitter dans le traitement de la campagne.

Le premier constat est simple : **le sujet des élections municipales mobilise la twittosphère**. Les chiffres le montrent : dans le cadre de notre étude, nous avons recueilli plus de 500 000 tweets émis entre le 1^{er} janvier et le 17 février 2020, ce qui correspond à près de 18 millions de mots émanant de quelque 158 000 twittos différents ! Et ce n'est pas fini : nous reviendrons dans quelques jours avec un nouveau rapport analysant les évolutions significatives observées dans cette twittosphère des Municipales, sur la deuxième partie de la campagne.

Bonne lecture

Emmanuel Rivière, Directeur Général, Division Public de Kantar

Jean Laloux, Directeur associé, Inférences

Christian de La Guéronnière, Directeur, Epiceum

CHIFFRES CLÉS



Données essentielles

528 000

tweets

(min-max par ville : de 6,8 à 302 K)

158 000

twittos

(min-max par ville : de 3,6 à 87 K)

18 000 000

mots

(min-max par ville : de 233 à 1055 K)

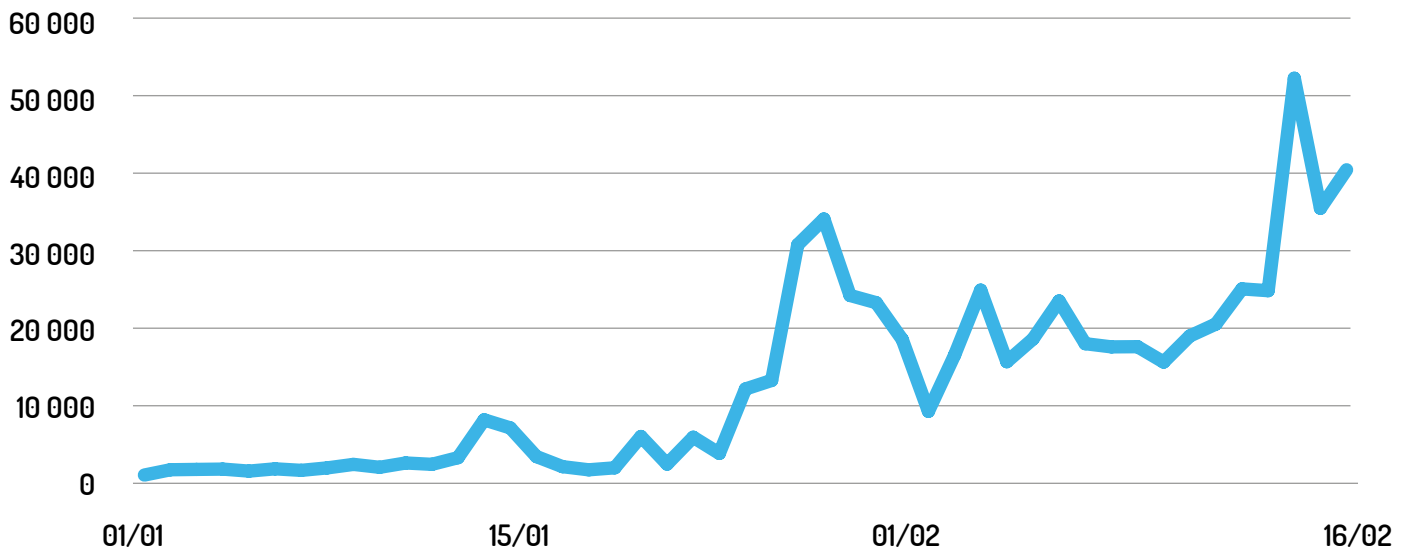
11

villes

23

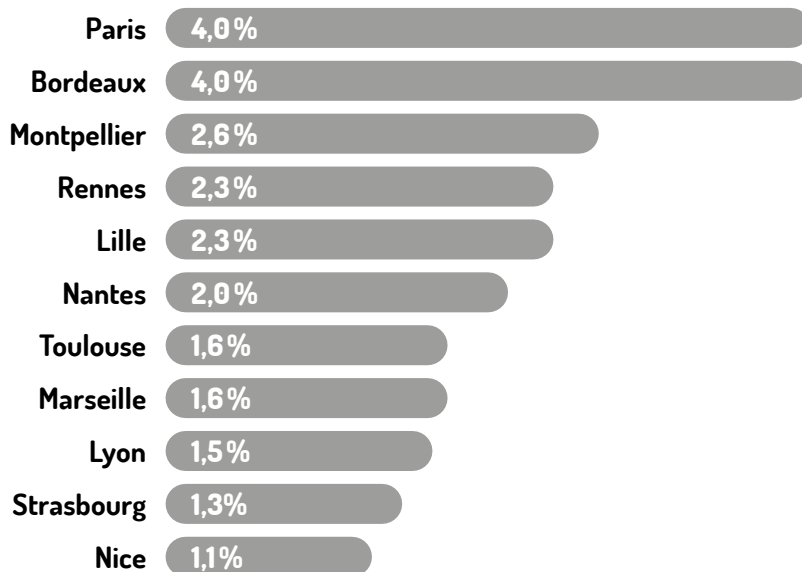
thèmes

Évolution du nombre de tweets



	Nb de tweets	Nb d'auteurs uniques	Nb de mots
Bordeaux	31 449	10 044	1 034 679
Lille	22 297	5 273	794 204
Lyon	22 010	7 691	726 649
Marseille	48 989	13 941	1 693 588
Montpellier	29 783	7 414	1 069 054
Nantes	20 049	6 064	680 595
Nice	12 853	3 873	455 665
Paris	302 005	87 623	10 555 254
Rennes	10 394	4 979	365 708
Strasbourg	6 812	3 622	233 156
Toulouse	21 636	7 693	732 833
TOTAL	528 277	158 217	18 341 385

Ratio auteurs uniques / habitants (en %)



Part du nombre d'auteurs de tweets sur chaque ville relative au nombre d'habitants pour chaque ville.

Ce graphique montre le ratio en % entre le nombre d'auteurs uniques dans chaque twittosphère et la population totale de la ville concernée. Paris et Bordeaux sont les villes qui ont les plus fortes densités de twittos s'exprimant sur elles, Strasbourg et Nice celles qui en ont le moins.

Nota : l'auteur unique citant une ville ne signifie pas forcément qu'il réside dans la ville (par exemple, ce peut être un tweet d'un média parisien qui enquête sur une ville en région, ou d'un média régional qui commente la campagne à Paris). Néanmoins, au regard du modèle d'extraction qui a centré les tweets sur des thèmes municipaux, il est raisonnable de penser que les auteurs uniques ont une forte probabilité d'être attachés d'une manière ou d'une autre à la ville concernée.

PRINCIPAUX ENSEIGNEMENTS

Municipales 2020 sur Twitter : les transports et l'écologie dominent les échanges sur Twitter. Paris imprime sa prééminence. Les partis politiques sont peu mentionnés.

De quoi parle-t-on au niveau national sur Twitter ?

Du côté des thèmes leaders (10% et plus des mentions), on observe une percée des sujets « verts » : l'**environnement** et le **transport**. Ce constat nous enseigne que ces sujets sont désormais transversaux, et non plus uniquement associés aux partis écologistes. Les nombreuses enquêtes d'opinion montrent en effet une montée de ces préoccupations en France et les maires disposant de capacités d'actions sur le cadre de vie et l'organisation de la ville, expliquent qu'elles trouvent naturellement leur déclinaison dans les débats locaux.

Un autre signal assez net concerne les sujets en lien avec la **sécurité** et les **forces de l'ordre** : la sûreté des biens et personnes est un sujet permanent de préoccupation, auquel s'ajoute la nouveauté de manifestations parfois violentes dans les centres-villes.

Concernant les thèmes challengers (5 à 10% des mentions), on retrouve la **culture**, la **démocratie participative**, les **projets** et l'**activité économique**.

Enfin, nous avons relevé des thèmes orphelins (moins de 5%). Si la **famille** et l'**enfant** sont présents dans les discours, les **seniors** sont les grands absents des conversations. Enfin, d'autres sujets sont très peu évoqués en ce début de campagne, comme la **santé** ou l'**emploi**.

Dans quelles villes tweete-t-on le plus ?

Paris et **Bordeaux** sont les villes qui ont les plus fortes densités de twittos s'exprimant sur elles ; **Strasbourg** et **Nice** celles qui en ont le moins.

La ville de Paris représente à elle seule 55% des twittos et 57% des tweets, révélant que l'usage du réseau social reste inégal au plan territorial et reflète les déséquilibres souvent constatés entre Paris et le reste du territoire, ce qui suggère que ce réseau pourrait contribuer à entretenir un prisme parisien dominant dans la lecture de l'actualité.

Classement des partis les plus mentionnés sur la twittosphère

Si les partis politiques sont peu mentionnés dans la twittosphère relative à la campagne des municipales à Paris et dans les 10 plus grandes villes de France, LREM se hisse sans conteste à la première place du classement avec plus de la moitié des mentions observées (53,6%). Les autres partis sont loin derrière. On note que le RN se place en 5^e position (5,3%), marquant un décalage avec sa place dans la vie politique nationale. En revanche, d'autres partis tels que LR (12,5%), EELV (9,9%) ou le PS (8,1%) ont un taux de citation plus conforme à ce qui est attendu par rapport à leurs scores récents.

FOCUS SUR LES THÉMATIQUES DE LA CAMPAGNE

Les tweets en lien avec les élections municipales ont été analysés de manière à faire ressortir les thématiques les plus abordées pour Paris et les 10 plus grandes villes des France.

LES MOTS DE LA PREMIÈRE PARTIE DE LA CAMPAGNE

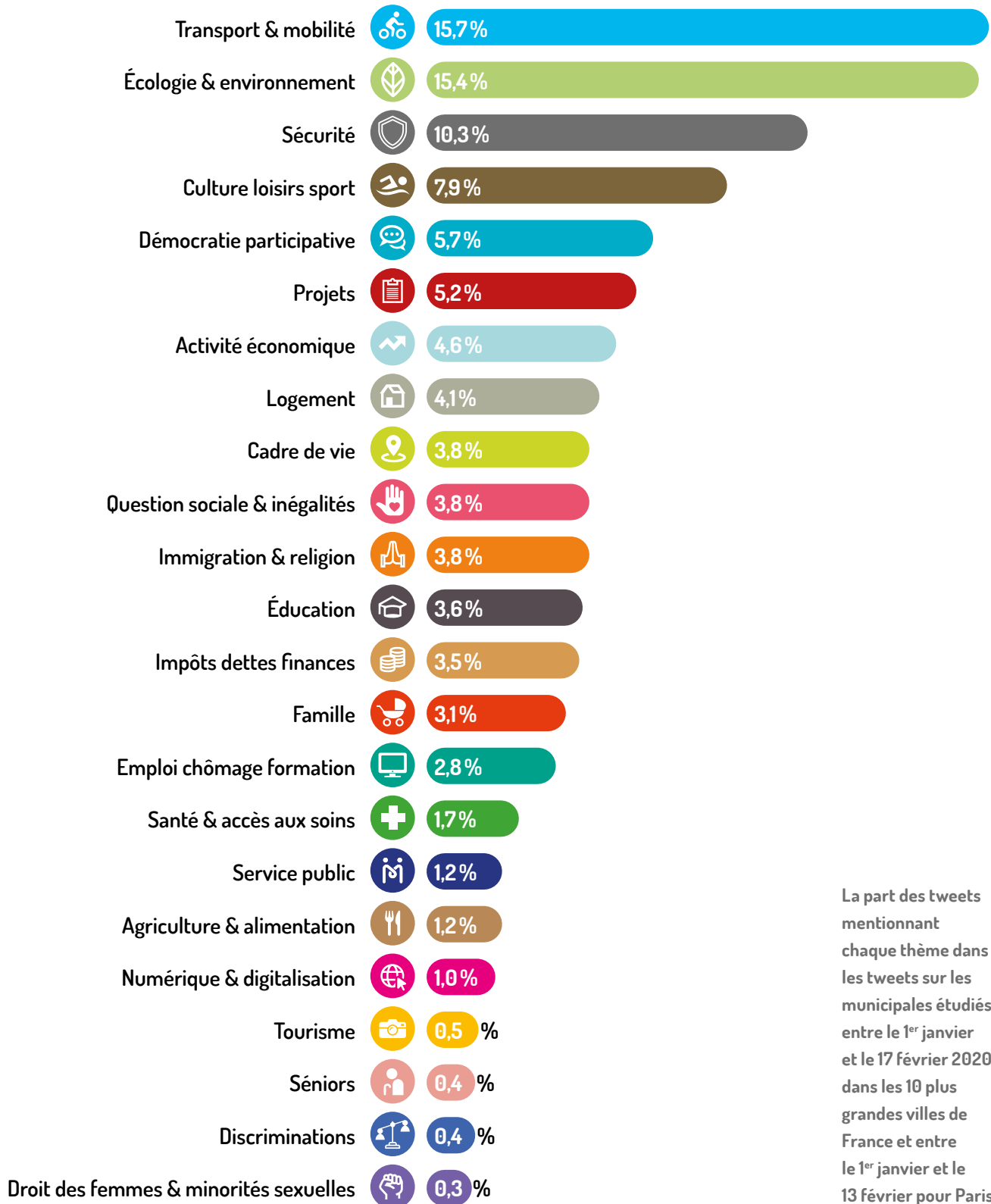


La taille des mots est proportionnelle à leur nombre d'apparition dans les tweets sur les municipales étudiés entre le 1^{er} janvier et le 17 février 2020 dans les 10 plus grandes villes de France et entre le 1^{er} janvier et le 13 février pour Paris.

Les projets, l'écologie, le vélo, le logement, la sécurité, la propreté, les citoyens... tels sont quelques-uns des topiques qui dominent le média Twitter lors des 45 premiers jours de la campagne municipale 2020. La percée sur le réseau des sujets « verts » concernant l'environnement, le cadre de vie et le transport est un premier enseignement. Ces sujets sont désormais transversaux, et non plus uniquement associés aux partis écologistes. Les électeurs préféreront-ils l'original à la copie, comme on le dit parfois ? C'est une des questions de ce scrutin. Un autre signal assez net concerne les topiques en lien à la sécurité et aux forces de l'ordre : la sûreté des biens et personnes est un sujet permanent de préoccupation, auquel s'ajoute la nouveauté de manifestations parfois violentes dans les centres-villes. Si la famille et l'enfant sont présents dans les discours, les seniors sont les grands absents des conversations. Et d'autres sujets sont très peu évoqués en ce début de campagne, comme la santé ou l'emploi.



LES THÈMES DE LA PREMIÈRE PARTIE DE LA CAMPAGNE



La part des tweets mentionnant chaque thème dans les tweets sur les municipales étudiés entre le 1^{er} janvier et le 17 février 2020 dans les 10 plus grandes villes de France et entre le 1^{er} janvier et le 13 février pour Paris.

Thèmes leaders (10% et plus des mentions) : les transports et l'écologie sont les deux thématiques qui dominent nettement les échanges sur Twitter, dans toutes les villes ou presque, avec en 3^e position la sécurité en léger décrochage.

Thèmes challengers (5 à 10% des mentions) : la culture, la démocratie participative, les projets et l'activité économique sont présents dans les échanges.

Thèmes orphelins (moins de 5%) : ce sont les oubliés de ce début de campagne. Certains peuvent être localement plus forts dans une ville que dans une autre, comme le logement, le cadre de vie, la question sociale ou l'emploi. D'autres sont franchement négligés partout comme la santé, les seniors, les discriminations, le numérique.

La domination des thèmes écologie et transport ne vient pas comme une surprise : les enquêtes d'opinion montrent une montée de ces préoccupations en France et les maires disposent de capacités d'action sur le cadre de vie et l'organisation de la ville.

Le résultat est cependant surprenant pour les thèmes négligés, car lorsque les enquêtes traditionnelles interrogent les Français, des sujets comme la santé, l'emploi, le logement arrivent très haut dans les réponses, et souvent avant l'environnement. On l'avait aussi observé dans le Grand Débat national de 2019. Or, les maires peuvent agir sur ces sujets.

Cela accrédite l'hypothèse d'une divergence entre l'expression des influenceurs sur Twitter et l'expression des citoyens pris dans leur ensemble.

CLASSEMENT DES THÉMATIQUES D'IMPORTANCE PRIORITAIRE POUR LES FRANÇAIS



Préoccupations citées comme prioritaire lors d'un sondage Kantar mené auprès des individus habitant des agglomérations de plus de 100 000 habitants en France.

Une comparaison entre les sujets prioritaires perçus par les habitants des grandes villes (100 000 habitants ou plus) pour l'avenir de leur ville et les sujets évoqués spontanément sur Twitter concernant les élections municipales dans les principales villes françaises permet d'observer des différences importantes, mais aussi certains points communs.

Si les transports et l'environnement se distinguent clairement en thèmes leaders sur Twitter, ces deux thèmes n'apparaissent cependant pas comme les plus prioritaires aux yeux des habitants des grandes villes (ils ne sont cités respectivement qu'à la 8ème et 5ème place). Dans les deux cas, une majorité des personnes interrogées les considèrent comme prioritaires, mais ils ne sont pas les sujets faisant le plus l'unanimité sur ce point. La priorité d'action pour les habitants des grandes villes est réservée aux sujets pérennes de la santé et l'éducation, et la troisième place, la sécurité. Cette différence dans le classement des sujets sur Twitter par rapport à leur ordre de priorité pour les citoyens indique que les discussions sur Twitter ont plus tendance à être tirées par des sujets d'actualité.

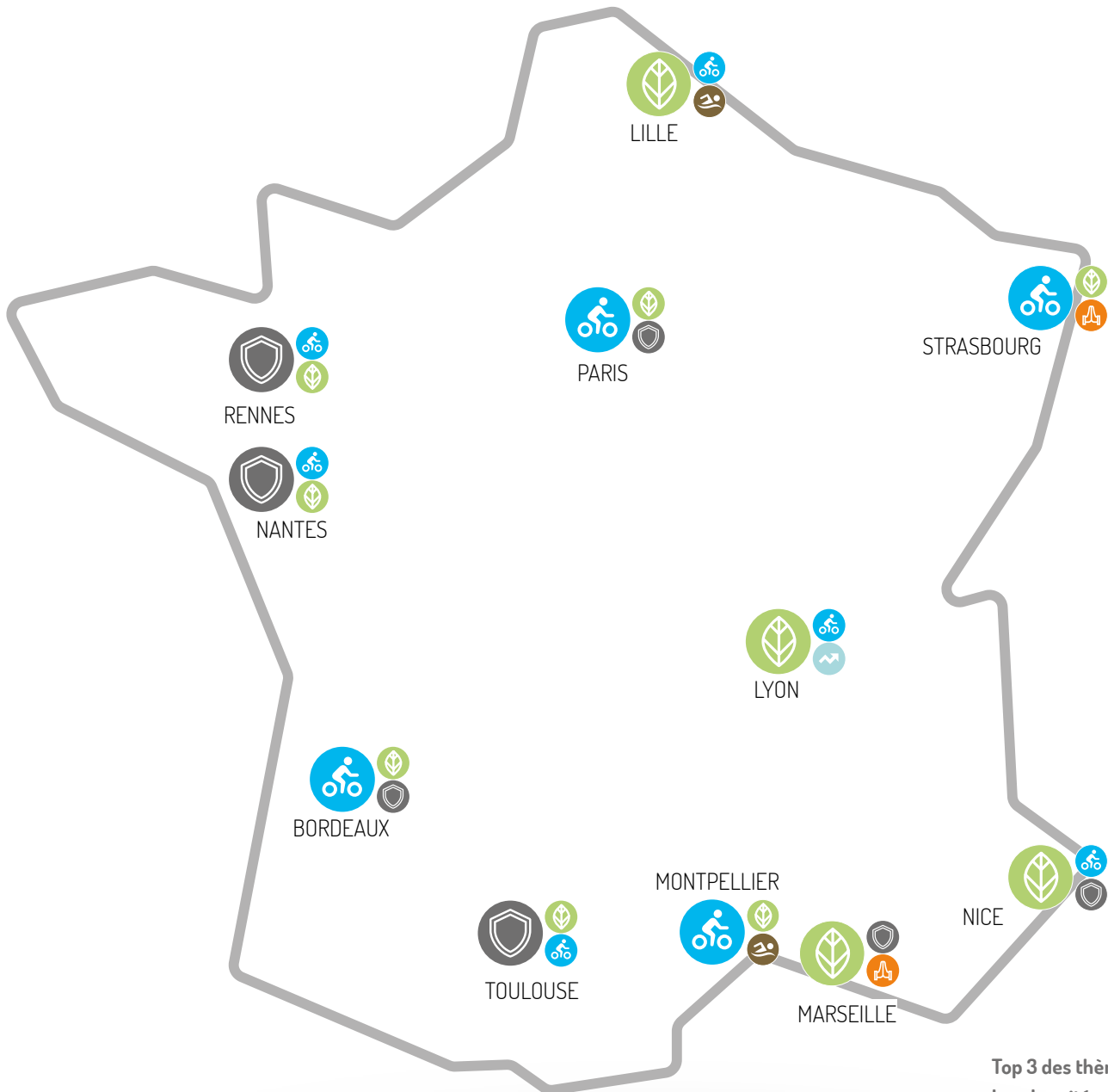
Concernant les transports, il est notable d'observer que les tweets évoquant les municipales à Paris sont de loin les plus nombreux (57% du corpus), et évoquent particulièrement cette question. La période de recueil des tweets du 1er janvier au 17 février 2020, dans une période particulièrement marquée par les grèves dans les transports parisiens, a probablement donné plus l'importance à ce sujet dans les discussions.

Un autre point de différence notable : la démocratie participative est le cinquième sujet le plus évoqué spontanément sur Twitter, et semble donc tenir une place relativement importante, tandis qu'il est cité comme le sujet le moins prioritaire pour les habitants des grandes villes, révélant un discours sur Twitter qui a plus tendance à parler des sujets qui suscitent de l'intérêt et de la réaction.







Si des différences importantes sont observables, certains points communs apparaissent aussi. La sécurité se distingue notamment comme un thème important dans les deux cas. Il est le troisième sujet le plus prioritaire aux yeux des habitants des grandes villes et le troisième sujet le plus évoqué spontanément sur Twitter, indiquant une discussion sur Twitter qui reconnaît l'importance de cette préoccupation.

Un autre point commun peut être observé : les sujets liés aux questions économiques et sociales ne sont pas les plus cités ni les plus prioritaires, même s'ils demeurent importants. Ce point commun démontre que ces thèmes ne seront sans doute pas les plus déterminants dans la campagne des municipales.

LES THÈMES LES PLUS CITÉS DANS LES 11 VILLES



Top 3 des thèmes les plus cités dans chaque ville, dans les tweets en lien avec les municipales 2020 entre le 1^{er} janvier et le 17 février 2020 pour les 10 plus grandes villes de France et entre le 1^{er} janvier et le 13 février 2020 pour Paris.

-  Sécurité
-  Écologie & environnement
-  Transport & mobilité
-  Activité économique
-  Immigration & religion
-  Culture loisirs sport

RÉSULTATS DES THÈMES PAR VILLE

	Bordeaux	Lille	Lyon	Marseille	Montpellier	Nantes	Nice	Paris	Rennes	Strasbourg	Toulouse
activité économique	6,8	5,1	6,2	3,1	4	6,8	3,6	4,4	2,7	3,8	3,8
agriculture & alimentation	1,7	1,2	1,0	1	1,2	1,7	0,7	1,0	2,2	0,7	1,0
cadre de vie	2,6	3,8	1,6	6,2	4,5	3,5	6,2	8,0	1,8	1,2	2,9
culture loisirs sport	7,7	7,3	5,9	5,9	12,7	7,5	10,0	6,5	11,0	5,9	6,8
démocratie participative	5,2	5,7	3,9	7,4	9,5	4,7	4,2	2,9	7,4	5,3	6,6
discriminations	0,2	0,9	0,2	0,2	0,1	0,2	1,0	0,4	0,2	0,6	0,3
droit des femmes & minorités s	0,4	0,5	0,4	0	0,1	0,0	0,2	0,6	0,1	0,3	0,1
écologie & environnement	14,1	18,4	25,5	14,2	13,7	13,9	16,4	11,3	12,6	14,2	14,6
éducation	2,3	2,9	4,3	4,2	3,2	5,5	2,7	1,6	4,4	4,7	4,1
emploi chômage formation	3,5	4,7	3,2	2,5	1,4	1,7	2,0	2,3	4,2	3,3	2,2
famille	3,7	2,3	2,4	5,2	6	1,9	4,1	4,1	2,0	0,5	1,9
immigration & religion	0,1	4,2	2,3	9,2	0,2	5,2	0,8	1,5	5,0	7,2	6,1
impôts dettes finances	2,4	6,6	2,2	2,1	1,9	1,7	6,0	2,0	2,9	6,9	3,5
logement	5,9	2,7	3,9	6,8	2,5	4,2	2,2	6,9	4,0	2,4	3,5
numérique & digitalisation	1,3	1,1	0,9	1	0,6	0,5	0,6	0,5	1,5	1,9	0,8
projets	8,1	4,4	4,2	3,3	4,2	3,5	8,7	8,9	3,0	4,8	4,5
question sociale & inégalités	3,4	5,6	3,4	3,1	7,5	2,2	1,9	4,3	2,6	5,5	2,7
santé & accès aux soins	1,2	1,4	1,2	2,6	1,9	3,6	1,5	1,9	1,5	1,0	0,9
sécurité	9,8	6,0	3,5	10,8	5,2	15,9	11,8	8,9	18,3	5,8	17,7
séniors	0,8	0,4	0,0	0,2	1,6	0,0	0,2	0,2	0,3	0,2	0,8
service public	1,6	1,3	1,7	1,1	0,7	0,9	0,8	1,5	1,4	1,5	1,2
tourisme	1,2	0,6	0,1	0,2	0,4	0,1	1,0	1,1	0,1	0,8	0,1
transport & mobilité	15,9	13,0	22,0	9	16,6	14,9	13,4	19,2	13,1	21,4	13,9

Ce tableau restitue le poids exact de chacune des 23 thématiques dans les 11 villes.

La couleur va du rouge vif (très forte citation) à bleu sombre (très faible citation) avec un gradient.

Constat 1 : sur la twittosphère, il y a une certaine **homogénéité nationale** de la campagne dans les grandes villes. On constate des lignes à dominante rouge uniforme partout (écologie, transport, culture) et d'autres à dominante bleu uniforme partout aussi (discrimination, droit des femmes, séniors, services public, tourisme, numérique, santé).

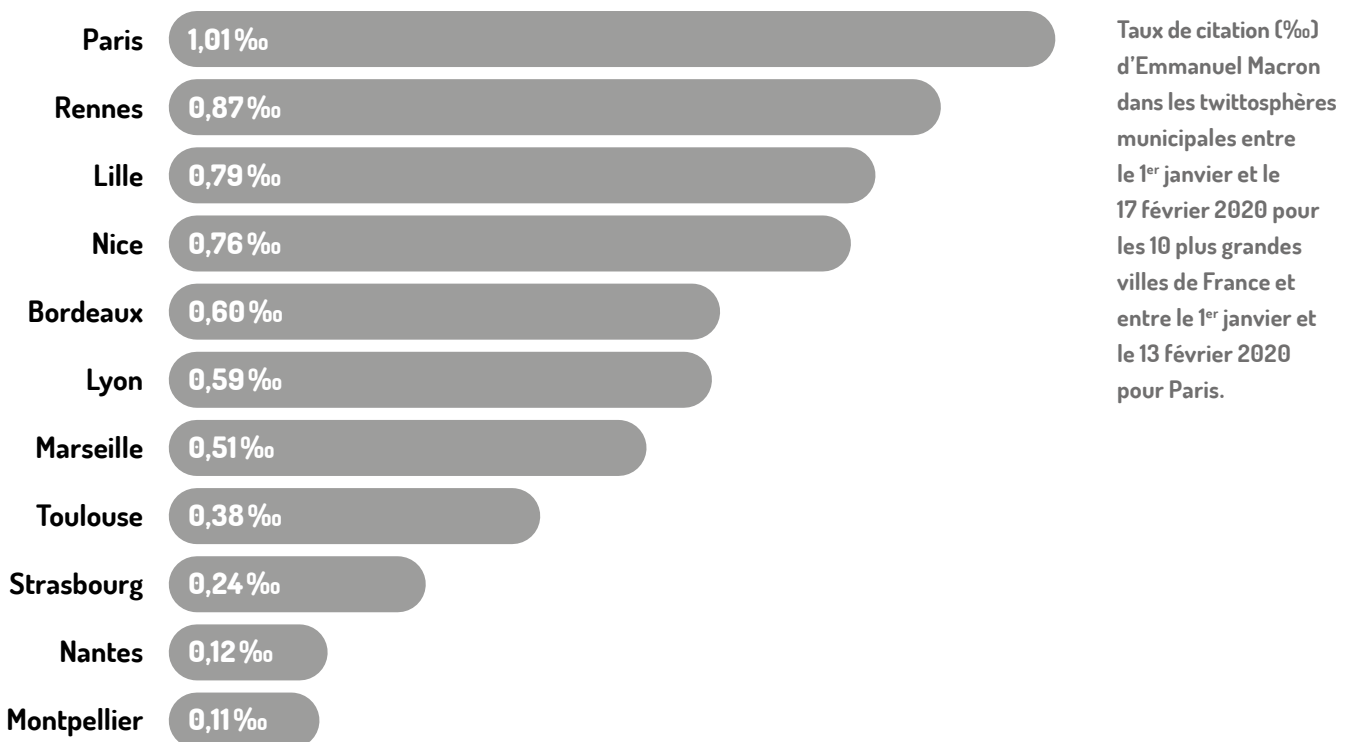
Constat 2 : certains thèmes marquent davantage de variation d'une twittosphère à l'autre (facteur 2 à 3 du taux de citation), sans qu'ils soient généralement les thèmes leaders. Ce sont plutôt des **signaux faibles de divergence** d'une ville à l'autre, par exemple sur la sécurité, l'immigration, la question sociale, le cadre de vie.

DES ÉLECTIONS INTERMÉDIAIRES POUR UN SCRUTIN NATIONALISÉ ? LA RÉFÉRENCE À EMMANUEL MACRON DANS LES ÉCHANGES

La question se pose à chaque scrutin local : s'agit-il d'une « élection intermédiaire » qui doit être interprétée comme un avis politique sur le gouvernement en place ? Et donc d'un scrutin « nationalisé » ?

Il n'est pas dans la portée de cette étude de déchiffrer avec certitude les intentions des citoyens postant sur Twitter. L'analyse des thèmes ne laisse toutefois pas présager un scrutin très national pour les 11 grandes villes : la question sociale est par exemple en retrait dans beaucoup de villes, tandis que les sujets purement locaux comme les transports mobilisent très fortement l'attention.

Nous avons mesuré le taux de citation d'Emmanuel Macron dans les échanges, pour vérifier si les twittosphères sont uniformes ou non.



On voit des différences significatives, puisqu'il existe un facteur 10 entre les twittosphères des villes citant le plus le président (Paris) et celles le citant le moins (Montpellier).

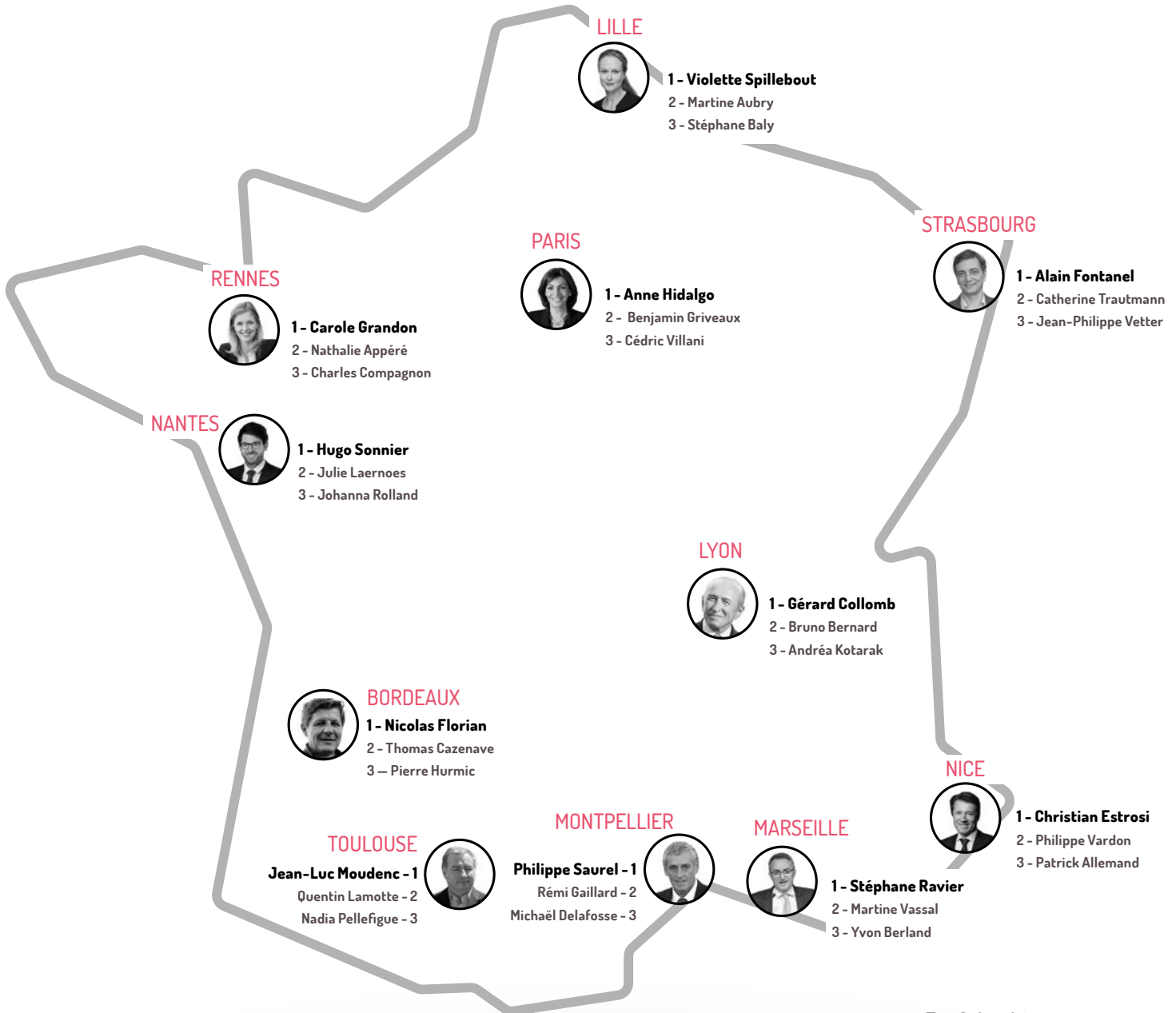
Mais ces taux de citation (en pour mille) restent très faibles, du même ordre que ceux (non représentés) sur la réforme des retraites.

On peut conclure que dans l'expression des twittos, l'élection municipale n'est globalement pas perçue comme un scrutin de « réponse à Macron », même si la capitale n'est pas exempte de cet arrière-plan pour une partie des citoyens postant sur le réseau.

FOCUS SUR LE PODIUM DES TÊTES DE LISTE

Le Twittorama analyse la manière dont la twittosphère évoque les têtes de liste les plus mentionnées.

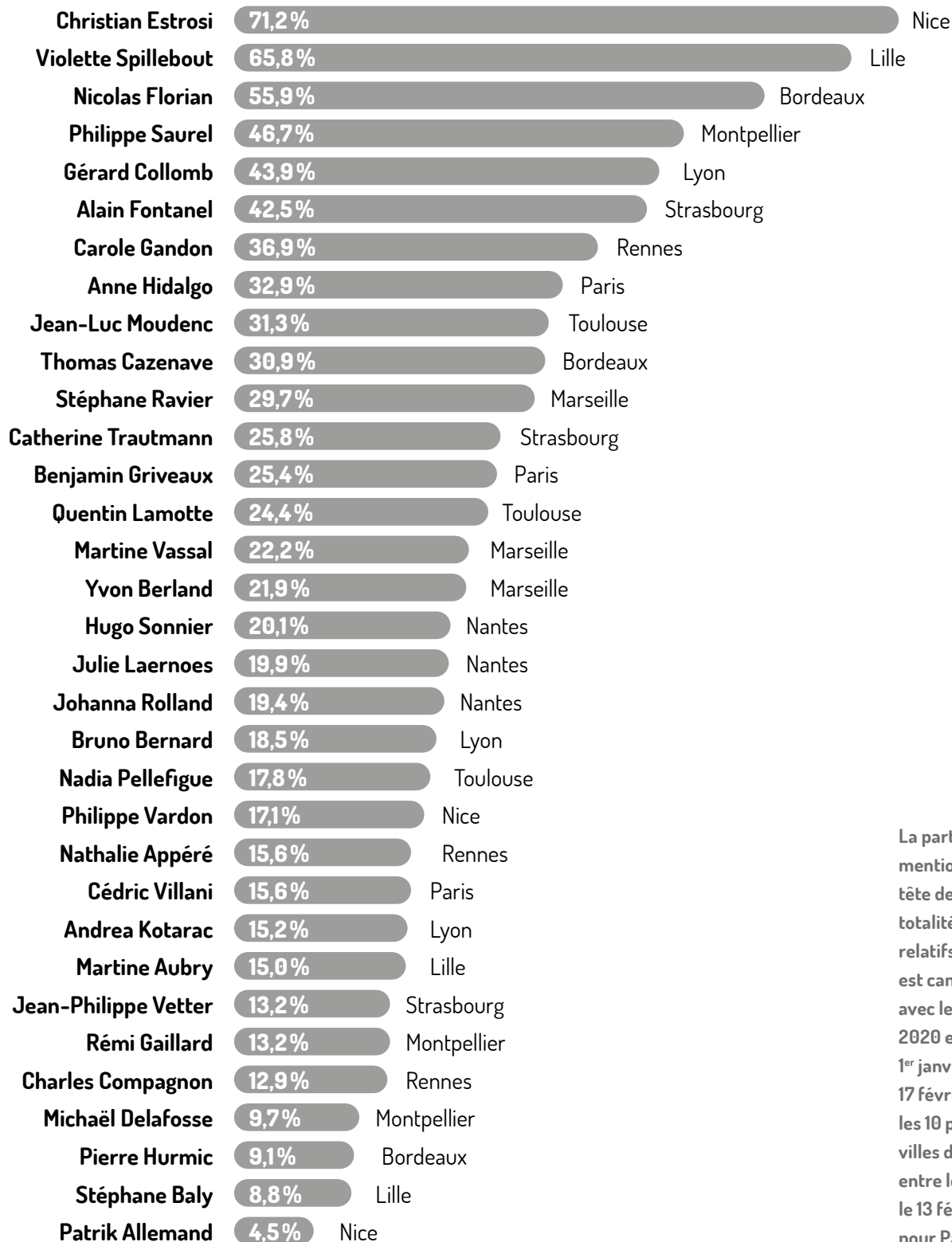
LES TÊTES DE LISTE LES PLUS CITÉES DANS LES 11 VILLES



Top 3 des thèmes les plus cités dans chaque ville, dans les tweets en lien avec les municipales 2020 entre le 1^{er} janvier et le 17 février 2020 pour les 10 plus grandes villes de France et entre le 1^{er} janvier et le 13 février 2020 pour Paris.



LE TOP 30 DES CANDIDATS LES PLUS CITÉS DANS LEURS VILLES



La part des tweets mentionnant chaque tête de liste dans la totalité des tweets relatifs à la où elle est candidate, en lien avec les municipales 2020 entre le 1^{er} janvier et le 17 février 2020 pour les 10 plus grandes villes de France et entre le 1^{er} janvier et le 13 février 2020. pour Paris.

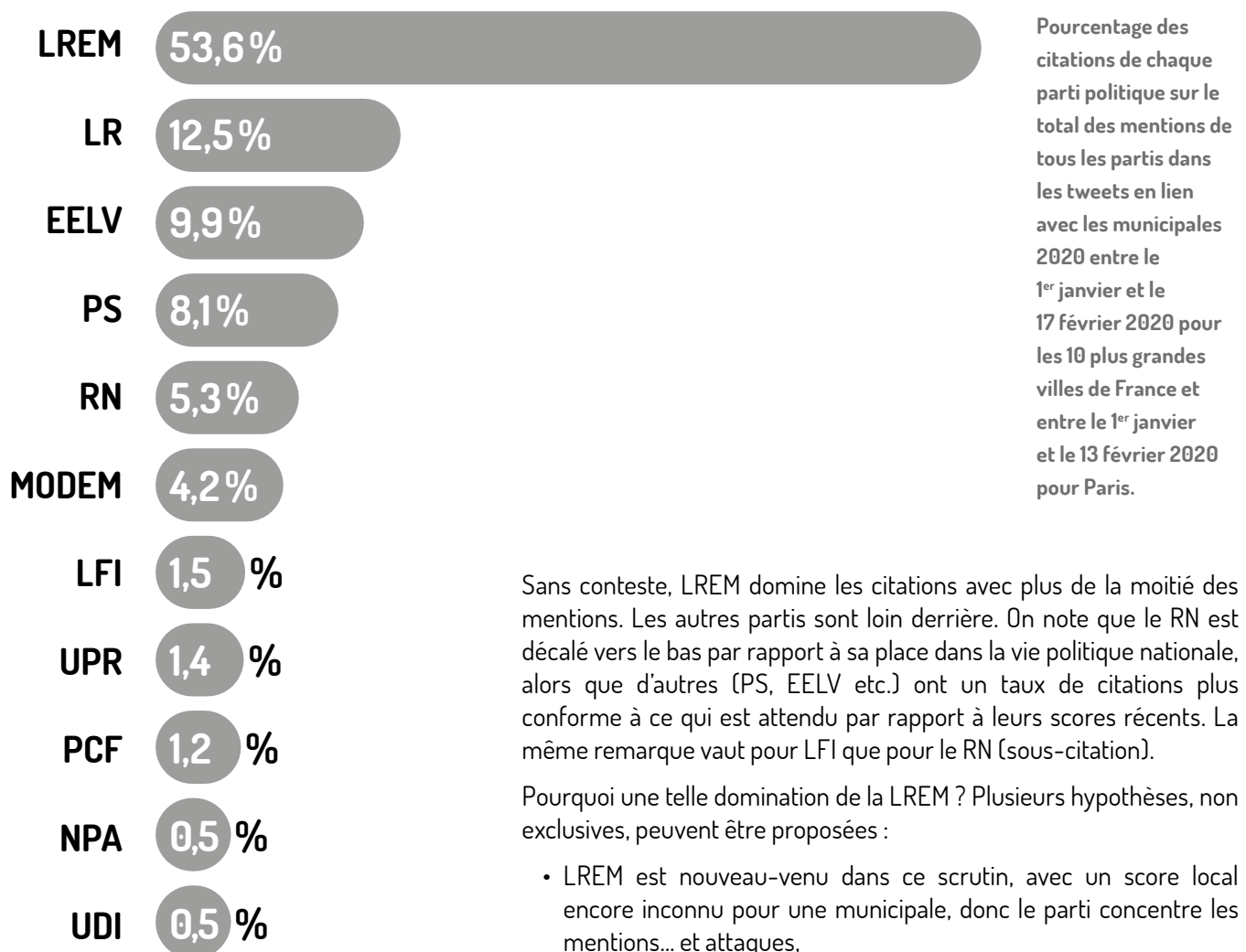
Si la valeur moyenne s'établit à 28,5%, on constate une forte dispersion des mentions de candidats sur les podiums : de 4,5 à 71,2%.

LES CITATIONS DES PARTIS POLITIQUES

Nous avons totalisé 89 882 citations de partis, étant entendu que cela ne comptabilise pas les noms de listes locales, mais bien les références aux sigles ou noms des structures partisanes nationales.

Ce chiffre d'environ 90 K est à rapporter aux plus de 18 millions de mots du corpus analysé : il ne représente que 0,49% des mots. Les hashtags des partis sont donc loin d'accompagner tous les tweets, ce qui corrobore d'autres observations faites sur la mise en retrait des structures partisanes dans ce scrutin.

Quels partis sont les plus cités ? Voici le taux de citation toutes villes confondues des dix premiers partis :



Sans conteste, LREM domine les citations avec plus de la moitié des mentions. Les autres partis sont loin derrière. On note que le RN est décalé vers le bas par rapport à sa place dans la vie politique nationale, alors que d'autres (PS, EELV etc.) ont un taux de citations plus conforme à ce qui est attendu par rapport à leurs scores récents. La même remarque vaut pour LFI que pour le RN (sous-citation).

Pourquoi une telle domination de la LREM ? Plusieurs hypothèses, non exclusives, peuvent être proposées :

- LREM est nouveau-venu dans ce scrutin, avec un score local encore inconnu pour une municipale, donc le parti concentre les mentions... et attaques,
- LREM a encore peu de militants locaux, ses cadres et relais utilisent davantage les médias sociaux,
- LREM est le parti au pouvoir, donc des faits politiques nationaux (plus souvent associé à son nom) peuvent être cités dans les échanges.



Taux de citation des partis par ville :

	Bordeaux	Lille	Lyon	Marseille	Montpellier	Nantes	Nice	Paris	Rennes	Strasbourg	Toulouse
cent (udi)	0,0	0,0	0,6	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0
cds	0,0	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0
cnip	0,0	0,0	0,0	0,0	0,0	0,0	0,9	0,0	0,0	0,0	0,0
dlf	0,0	0,0	0,0	0,0	0,1	0,0	4,6	0,0	0,0	0,0	0,1
eelv	5,5	10,4	19,1	4,6	35,5	6,9	3,3	9,4	24,7	7,1	1,8
ensemble	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
ge	0,0	0,0	0,1	0,1	0,0	0,1	0,0	0,0	0,0	0,0	0,0
lfi	2,6	2,3	2,0	1,0	0,6	0,5	1,2	1,5	0,5	1,6	0,7
lo	0,0	0,1	0,9	0,0	0,0	1,4	0,3	0,0	0,1	0,0	0,1
lp	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
lr	8,3	6,9	13,0	12,2	2,5	3,0	13,6	18,4	2,6	3,1	7,0
lrem	59,3	59,9	29,8	72,5	41,4	33,9	44,7	53,0	33,3	37,4	45,0
mdp	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0
mei	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0
modem	4,1	6,8	2,8	2,0	4,7	1,8	1,2	5,5	4,7	2,0	2,1
mrc	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,5	0,0	0,0	0,2
nc	0,0	0,1	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,0
npa	5,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2
pcd	0,0	0,0	0,8	0,0	0,0	0,0	0,0	0,3	0,0	0,0	0,4
pcf	2,7	0,4	1,5	0,2	0,5	0,4	1,1	1,5	0,2	2,4	0,9
poi	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0
prg	0,2	0,6	0,2	1,3	0,0	0,0	0,2	0,4	0,0	0,0	0,0
ps	4,1	10,1	6,7	5,6	6,0	22,9	6,8	6,3	21,3	42,4	2,7
rn	1,1	1,5	19,8	0,0	7,4	9,1	20,5	2,6	9,0	3,2	38,7
ude	0,2	0,2	0,4	0,1	0,0	0,1	0,0	0,0	0,0	0,0	0,0
udf	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0
udi	1,2	0,5	0,5	0,1	0,8	0,2	0,0	0,5	2,9	0,0	0,2
upr	5,5	0,0	1,8	0,0	0,2	19,5	1,5	0,1	0,3	0,9	0,0

Les nuances de couleur permettent de voir rapidement les partis très cités (vers le rouge) ou très peu cités (vers le bleu) dans chaque ville.

Si LREM est uniformément dominant des mentions dans toutes les villes (Marseille en tête, Lyon en queue), les autres partis peuvent connaître davantage de variations : par exemple forte présence EELV à Montpellier et Rennes, citation soutenue du PS à Strasbourg, Rennes et Nantes, mention fréquente du RN à Nice et Lyon. Et la surprise UPR à Nantes.

MÉTHODOLOGIE DE CETTE ÉTUDE

Formation du corpus

Le corpus analysé dans ce rapport est constitué de tweets émanant du réseau social Twitter discutant des élections municipales 2020 à Paris et dans les 10 plus grandes villes de France par l'importance de leur population. Il a été fait le choix d'intégrer les re-tweets dans ce corpus, car la popularité et la viralité d'un message font partie de l'écosystème d'information et d'influence de Twitter.

Les tweets ont été collectés et extraits via une requête booléenne de recherche homogène. Pour chaque ville, une requête de collecte a été constituée en associant le nom et les habitants de la ville (Rennes, Rennais, Rennaises, etc.) avec des mots-clés relatifs aux élections (élections, municipales, etc.) ou les noms ou comptes des candidats et des partis postulant dans la ville. Les hashtags associés aux campagnes des candidats dans chaque ville ont aussi été intégrés dans la requête de collecte.

Les tweets émis par les comptes des candidats eux-mêmes ou les comptes officiels de campagne ont été exclus afin d'éviter de surpondérer une « parole officielle ». Cependant, les re-tweets de messages émanant de ces sources sont intégrés pour prendre en compte l'influence que peuvent avoir ces retweets sur le discours global.

Une requête booléenne est toujours un compromis empirique entre un spectre trop large qui extrait des messages hors sujet et un spectre trop étroit susceptible de manquer des messages pertinents sur le sujet étudié. La requête mise en place a permis de recueillir plus de 500 000 tweets totalisant plus de 18 millions de mots et émanant de plus que 158 000 comptes uniques sur Twitter et a été vérifiée pour assurer un taux de pertinence supérieur à 95 %, constituant une base significative pour évaluer les tendances des échanges.

La date de collecte de ce premier corpus court du 1^{er} janvier au 17 février 2020 à 12h. Paris a formé une exception imprévue en raison de l'annonce le 15 février 2020 au matin du retrait de la candidature de Benjamin Griveaux. Cet événement a entraîné un afflux considérable de messages sans rapport avec notre prisme d'analyse de la campagne des Municipales 2020. Pour cette raison, le corpus de Paris a été arrêté au 13 février 2020 inclus. Un corpus complémentaire sur Paris a été formé du 17 au 21 février 2020 afin de disposer de premières informations sur la candidature d'Agnès Buzyn.

Classement quantitatif et matrice sémantique

Les scores de présence des candidats sont obtenus par la comptabilisation de leur patronyme et de leur nom de compte Twitter, le second étant le mode le plus fréquent de désignation sur le réseau social. Les mentions des candidats ne sont pas assimilables à un score d'intention de vote ou de popularité politique : il s'agit d'un indicateur de présence dans la twittosphère, pouvant être motivée pour des raisons positives, négatives ou neutres.

L'analyse des thématiques a été réalisée par des techniques de traitement semi-automatique du langage :

- Apprentissage supervisé à partir de plusieurs sources (corpus de textes politiques d'élus locaux, mots associés aux maires dans les médias - presse nationale et régionale - sur une année et enfin un mois de relevés de tweets à partir de la requête booléenne d'extraction précisée ci-dessus).
- Étiquetage et organisation raisonnée des données. Plus de 1 100 topiques (mots, syntagmes, locutions, sigles) sont venus nourrir 23 thématiques. La plupart de ces topiques sont des expressions à valeur univoque et générique (par exemple, « changement climatique » dans le thème « écologie & environnement »). Certains sont à valeur contextuelle et spécifique (par exemple « Central Park » et « gare de l'Est » dans le thème « projets » en référence au programme d'un candidat).
- La matrice sémantique ainsi constituée est donc une base d'interprétation globale, qui a suivi un apprentissage de perfectionnement et une mise en adéquation sur le corpus cible. Des tests aléatoires de contrôle ont été réalisés pour évaluer sa fiabilité en phase d'apprentissage puis lors des premières consolidations sur analyse des données globales réelles. On peut estimer la fiabilité d'attribution à plus de 95% de résultats corrects. Il convient de noter que cette matrice, conçue pour l'analyse des politiques publiques locales, ne quantifie pas les opinions individuelles, car ce n'est pas son objet. Par exemple, elle identifie comme relevant du thème « transports & mobilité » un tweet sur les « embouteillages », mais n'évalue pas si ce tweet est une louange ou un reproche ou un constat neutre. Un autre classifieur sémantique peut être utilisé à cette fin, mais l'analyse d'opinion ne formait pas l'objet de la présente étude. En traitement du langage, comme dans l'ensemble des techniques d'apprentissage supervisé souvent appelées improprement « Intelligence Artificielle », chaque problématique fait l'objet d'une modélisation *ad hoc*.



Les 23 grandes thématiques ont émergé de l'analyse sémantique (démarche inductive) et n'ont donc pas été définies a priori. Elles sont ici présentées par ordre alphabétique :

- | | | | |
|----|--|----|-------------------------------|
| 1 | Activité économique | 13 | Impôts dettes finances |
| 2 | Agriculture & alimentation | 14 | Logement |
| 3 | Cadre de vie | 15 | Numérique & digitalisation |
| 4 | Culture loisirs sport | 16 | Projets |
| 5 | Démocratie participative | 17 | Question sociale & inégalités |
| 6 | Discriminations | 18 | Santé & accès aux soins |
| 7 | Droit des femmes & minorités sexuelles | 19 | Sécurité |
| 8 | Écologie & environnement | 20 | Séniors |
| 9 | Éducation | 21 | Service public |
| 10 | Emploi chômage formation | 22 | Tourisme |
| 11 | Famille | 23 | Transport & mobilité |
| 12 | Immigration & religion | | |

Les 11 villes analysées par ordre décroissant du nombre d'habitants

- | | | | |
|---|-----------|----|-------------|
| 1 | Paris | 7 | Montpellier |
| 2 | Marseille | 8 | Strasbourg |
| 3 | Lyon | 9 | Bordeaux |
| 4 | Toulouse | 10 | Lille |
| 5 | Nice | 11 | Rennes |
| 6 | Nantes | | |

Résultats comparatifs du sondage

Les résultats d'un sondage représentatif de la population adulte des agglomérations françaises de 20 000 habitants ou plus sont présentés dans ce rapport pour donner un point de comparaison entre le poids des thématiques discutés sur Twitter et leur niveau de priorité pour les citoyens. Ce sondage a été réalisé auprès de 1 000 personnes du 14 au 17 février 2020, et les résultats présentés dans ce rapport ont été filtrés sur les habitants des agglomérations de 100 000 habitants ou plus pour correspondre aux populations des villes étudiées dans cette analyse, soit une base de 524 réponses.

Les résultats affichés présentent, pour chaque sujet, la proportion de réponses « Prioritaire » à la question suivante : « Pendant la campagne des élections municipales, les candidats vont débattre et faire des propositions autour de nombreux sujets. En pensant à l'avenir de votre ville pour les 6 prochaines années, veuillez indiquer pour chacun des sujets suivants s'il vous semble : Prioritaire / Important mais pas prioritaire / Secondaire / Sans intérêt »



Un outil au service de la démocratie participative

La Vox Populi est un outil de restitution et d'analyse de la parole des citoyens qui recourt aux technologies les plus récentes de traitement automatique de la langue couplées à une analyse humaine. L'objectif est de proposer une analyse sémantique fiable, indépendante et objective des contributions citoyennes recueillies dans les dispositifs participatifs ou dans des corpus textuels importants.

La Vox Populi est produite par l'expertise associée de deux entreprises :

- > **Inférences** : cabinet d'études expert en analyse sémantique, concepteur de la solution Opinion miner® d'analyse semi-automatique de grands volumes de données textuelles
- > **Epiceum** : agence conseil dédiée à la communication d'intérêt général, spécialisée dans l'accompagnement des politiques publiques et des grands projets, incluant notamment leurs périodes de débats publics.



Understand people, Inspire growth

Kantar est le leader mondial des études et du conseil fondés sur des données objectives. Nous délivrons une compréhension unique et exhaustive de la façon dont les individus pensent, ressentent et agissent, à l'échelle mondiale et locale, à travers plus de 90 marchés. En combinant l'expertise de nos équipes, nos bases de données, nos analyses et technologies innovantes, nous aidons nos clients à comprendre les individus et à stimuler leur croissance : Understand People, Inspire Growth.

En France, Kantar s'appuie sur l'expertise internationale et sur l'expérience éprouvée de la division Public pendant plus de 5 décennies, de ses équipes françaises (anciennement TNS Sofres) au service du secteur public, des organismes publics, des ONG, des universités et des grandes entreprises sur leurs enjeux de politiques publiques, de qualité de service public et de communication publique. Nous sommes en étroite relation avec les autres sociétés du groupe Kantar, pour bénéficier de la capacité de Kantar à innover et investir dans de nouveaux outils, et chercher en permanence à renouveler nos méthodes.

Retrouvez l'ensemble des résultats sur

lavoxpop.fr / fr.kantar.com



www.lavoxpop.fr
www.kantar.com



@LaVoxPop
@KantarFR



contact@lavoxpop.fr
info.fr@kantarpUBLIC.com