
Les Françaises et Français face à l'information

Max Chateau - Rasim Erben - Myndie Ferrandez - Quentin Jacquot

Polytech Montpellier - DaMS4



1 Méthodologie -

1.1 Enjeux et objectifs du projet -

Récemment, un article faisant la une du 13 Janvier du *Monde* titrait *Le Rassemblement national et ses idées s'installent dans l'opinion*. Toutefois comme le pointent certains sociologues comme Félicien Faury ou Vincent Tiberj, les médias, et en particulier la télévision, jouent un rôle essentiel dans la diffusion d'idée dans la société. En parallèle, plusieurs études et rapport montre un décalage entre la société française et ses médias, on peut évoquer le temps de parole des femmes qui en 2023 ne représentait uniquement 34% à la télé et la radio alors qu'elles représentent la moitié de la population.

De surcroît, le paysage audiovisuel français connaît de nombreuses transformations. La télévision et la radio, longtemps considérées comme les principales sources d'information, sont aujourd'hui concurrencées par les réseaux sociaux et les plateformes numériques. Une part importante des Françaises et des Français s'informe désormais en ligne, ce qui modifie les habitudes de consommation de l'information et peut influencer la confiance accordée aux médias traditionnels.

Ces constats soulèvent des questions importantes sur les biais présents dans l'offre d'information et sur leurs effets possibles sur le rapport des citoyens à l'actualité. Dans ce contexte, ce projet a pour objectif d'analyser les transformations du paysage audiovisuel français, en s'intéressant à la fois aux contenus diffusés (thèmes traités et représentation des femmes) et au rapport des Français à l'information (pratiques et confiance). La problématique centrale de ce travail est donc la suivante : dans quelle mesure les transformations du paysage audiovisuel français influencent-elles le rapport des Français à l'information, notamment en termes de pratiques et de confiance ?

Nous traiterons ainsi dans un premier temps de l'analyse thématique concernant les chaînes de TV, puis nous verrons

1.2 Les Datasets utilisés -

Les données du challenge

Sur le site du challenge, nous avons exploité les datasets suivants :

- ◆ Audience de la télévision - CNC
- ◆ Consommation VOD - CNC
- ◆ Télévision de ratrappage - CNC
- ◆ Classement thématique des journaux TV - INA

-
- ◆ Classement thématique des sujets de JT - INA
 - ◆ Temps de parole des hommes et des femmes à la télévision et à la radio - INA

Par manque de temps nous n'avons pas pu analyser en détail le jeu de données de l'ARCOM et en extraire le contenu concernant *Les français et l'information*, qui avait l'air de contenir des données intéressantes pour corroborer notre analyse, mais n'étaient pas aussi simple à exploiter que les autres jeux de données.

Le site data de l'INA

L'INA met à disposition de tout le monde le site data.ina.fr, ce dernier regroupe les analyse d'une IA sur les flux audios des principales chaines de la télévision. Les chaines d'information en continue présentes sont : BFM TV, CNews (anciennement I-Télé), France Info, et LCI. Et celles des chaines généralistes sont : TF1, France 2, France 3, M6 et Arte.

Différents axes sont présents pour chaque chaines de manière journalière :

- ◆ Proportion temps de parole Femmes/Hommes
- ◆ Nombre de fois qu'un mot a été prononcé
- ◆ Top 50 des personnes évoquées

Nota Bene : Ce données ne sont (en théorie) pas accessibles de manière brut publiquement, on a donc développé un petit programme pour communiquer avec l'API de l'INA et ainsi récupérer les données (dispo sur Github).

Jeux de données annexes ou non exploités

On a contacté Médiamétrie afin d'essayer d'obtenir des données sur les audiences plus précises, sans réponse de leur part malheureusement...

On s'est également basé sur l'actualité afin de pouvoir expliquer certaines données obtenues. On a également utilisé certains sondages lorsque cela nous semblait pertinent pour appuyer notre argumentaire.

1.3 Répartition du travail

Les différentes parties du rapport ont été rédigées par :

- ◆ Analyse du rapport des français aux médias : Quentin Jacquot
- ◆ Analyse thématique des chaines d'info en continu : Max Chateau
- ◆ Analyse thématique des chaines généralistes : Myndie Férrandez
- ◆ Présence des femmes dans les médias : Rasim Erben

La mise en forme globale du rapport a été réalisé par Max Chateau

Le site <https://datascience.axithem.fr> permettant d'observer les données traitées a été réalisé et déployé par Max Chateau. Quentin Jacquot a rajouté ses graphiques dans la partie audience.

2 Analyse thématique -

2.1 Chaines d'information en continu

Dataset utilisé : data.ina.fr

Les données n'étant pas classées par thème initialement, on a établit une liste d'une centaine de termes auxquels on a assigné un thème (ex : on a attribué le terme "Immigration" est associé au thème extrême droite). Cette liste est disponible dans la partie méthodologie de la section thématiques du site rendu. Les limites de cette méthode étant que les listes ont pu être mal établies d'une part (biais idéologique, nombre de termes insuffisants, etc...), d'autre part on n'a pas le contexte dans lequel le mot a été prononcé.

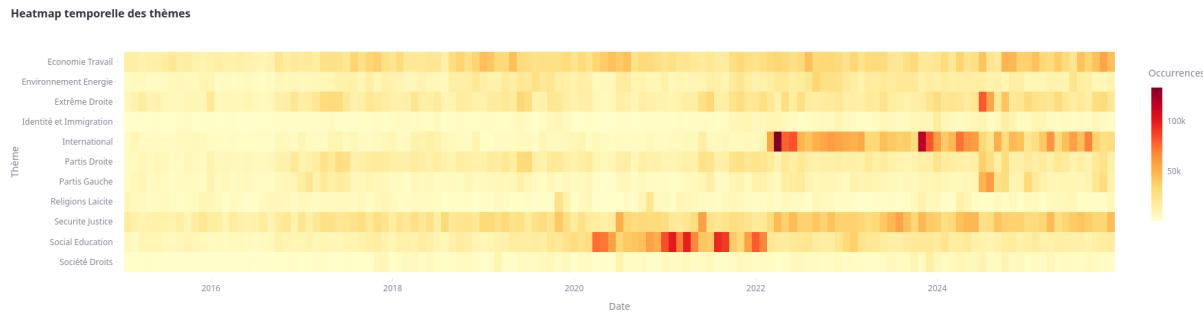


FIGURE 1 – Heatmap temporelle des thèmes sur les chaines d'info en continu

Evolution des thèmes diffusés

Nota Bene : Un thème "Vie institutionnelle" est présent également, mais les termes présents dedans sont tellement mentionné qu'ils éclipsent les autres thèmes, donc pour cette première analyse on fait fi de ce thème.

La vue globale nous montre plusieurs info intéressantes. Tout d'abord à partir de 2020 on constate une surreprésentation du thème social, notamment dû au fait que les termes liés au COVID sont compris dedans.

On constate également qu'à partir de Février 2022, le thème social est quasiment entièrement remplacé par les thèmes Sécurité/Défense et International, date du début de l'invasion Russe en Ukraine.

Un pic réapparaît pour ces mêmes thèmes en octobre 2023, date de l'attaque du Hamas en Israël, puis des attaques successives contre la Bande de Gaza.

Le dernier pic notable est en Juin 2024 concernant les thèmes liés à l'extrême droite, ainsi que la gauche dans une moindre mesure (on reviendra sur ce point plus tard), plus précisément lors du résultat des européennes et après l'annonce de la dissolution de l'assemblée nationale.

A l'inverse on constate que certains thèmes sont moins représentés que d'autres sur le temps long et n'ont que quelques représentation épisodique. On peut prendre l'exemple des sujets liés aux sujets société/droits et l'environnement. Bien que ces sujets semblent être important pour les français, comme le montre le Baromètre de l'IPSOS de septembre 2025. Par exemple, le sujet de l'environnement apparaît comme sous représenté, or il apparaît comme étant une préoccupation importante pour 20% des français en septembre 2025. Ceci peut être expliqué en partie par le fait que les chaines d'infos sont possédées par des entreprises privées, qui peuvent avoir des intérêts économiques contraires à la mise en avant de ces sujets.

Corrélation entre thèmes et chaines -

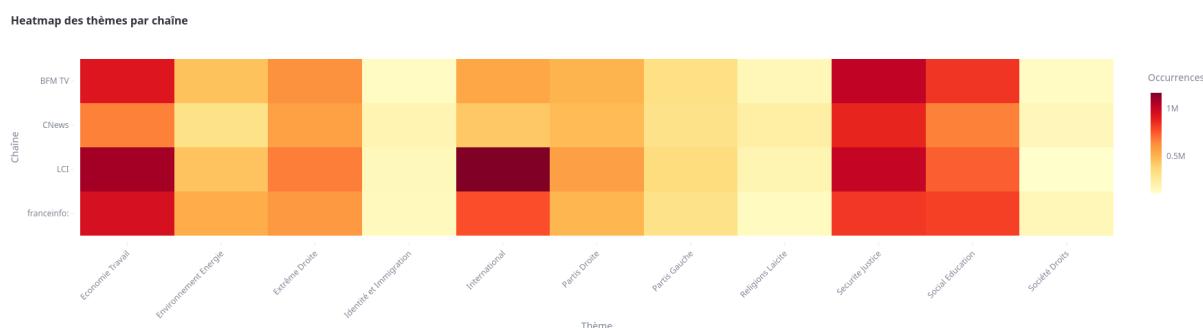


FIGURE 2 – Heatmap de la présence des thèmes sur les chaines d'info en continu, période 2015-2025

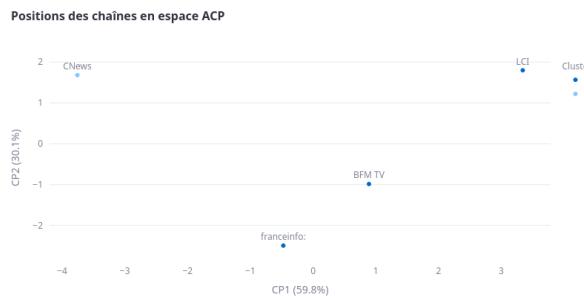


FIGURE 3 – ACP des chaînes d'information en continu sur les thèmes, période 2015-2025

influencer la ligne éditoriale de la chaîne.

On remarque également une sous représentation des thèmes économiques sur CNews par rapport aux autres chaînes depuis 2021, tandis que ce sujet est en augmentation pour les autres chaînes.

En se concentrant sur certains thèmes on constate une surreprésentation de la thématique de l'identité sur CNews par rapport aux autres chaînes. Cette tendance commence apparaître en 2018 (date où Vincent Bolloré reprend en main la chaîne), se calme en 2020 avec le COVID, puis repart en 2021. Tandis que dans le même temps les thèmes liés à l'environnement et aux droits/société sont moins représentés sur la chaîne. En appliquant une régression linéaire temporelle sur les thèmes de l'identité/immigration et de l'extrême droite, on obtient respectivement un $R^2_{id} = 0.48$ et $R^2_{ed} = 0.29$ avec une pente en hausse : $\beta_{1,id} = 24$ et $\beta_{1,ed} = 42$ tandis que sur les autres chaînes on constate également

On constate également que la Droite et l'extrême droite sont plus représentés que la Gauche sur l'ensemble des chaînes.

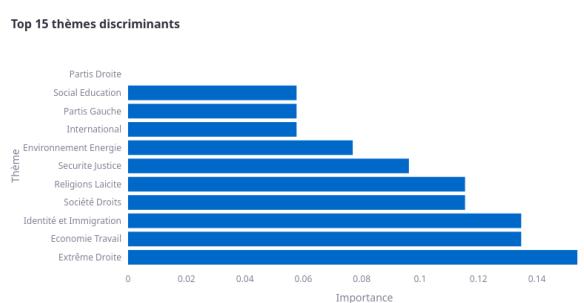


FIGURE 4 – Importance des thèmes pour la classification à l'aide d'une random forest sur toute la période

On constate en regardant la période entière disponible, que LCI a une surreprésentation des sujets internationaux, et se démarque également sur l'ACP, CNews se démarque également sur ses thèmes abordés, tandis que BFM TV et FranceInfo sont similaires. L'analyse des importances nous renseigne par le fait que les principaux marqueurs de séparation étant les thèmes liés aux partis de l'extrême droite, de l'économie et de l'identité. En modifiant les dates de l'ACP on constate qu'à partir de 2024, BFM se rapproche de LCI, on peut supposer que le rachat par Rodolphe Saadé a pu

On constate lors des échéances électorales que les thèmes politiques prennent une place plus importante sur les chaînes d'infos. Avec une très forte hausse lors des législatives de 2024, notamment au niveau de l'extrême droite notamment, suivi par la gauche et la droite dans une moindre mesure. En appliquant une régression sur le nombre d'évocation des partis de gauche et d'extrême droite, on constate, un forte corrélation entre celle-ci avec $R^2 = 0.71$ et un coefficient $\beta_1 = 0.7389$, marquant une surreprésentation de l'extrême droite par rapport à la gauche.

On ainsi peut re évoquer l'article cité en introduction, qui montrait la progression des idées du Rassemblement National dans l'opinion publique. On peut légitimement se demander si cette progression n'est pas en partie due à la surreprésentation de ces thèmes dans les médias audiovisuels, et notamment sur CNews.

L'influence des chaînes entre elles

On constate que les thèmes de Police/Justice ont considérablement augmenté sur la plupart des chaînes, mise à part BFM TV qui était déjà au dessus des autres (régression sur chacune des chaînes). Ces sujets étant en général associés aux faits divers, on peut se dire qu'après donné que BFM a été pendant longtemps la chaîne d'info la plus regardée en France, les autres chaînes ont pu s'inspirer de son succès en augmentant la part de ces sujets dans leurs journaux.

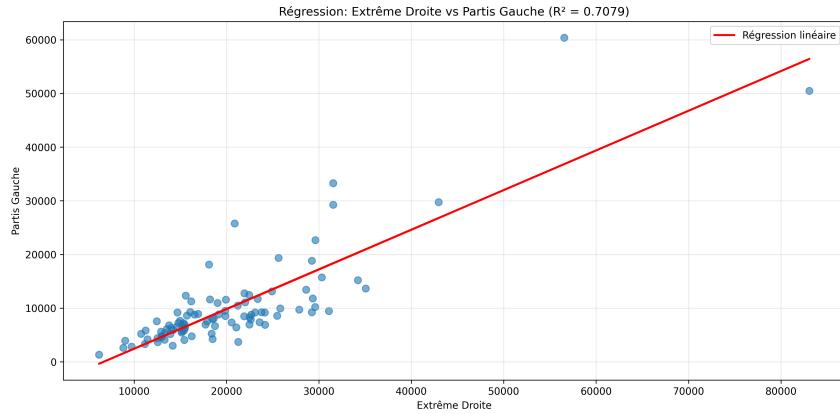
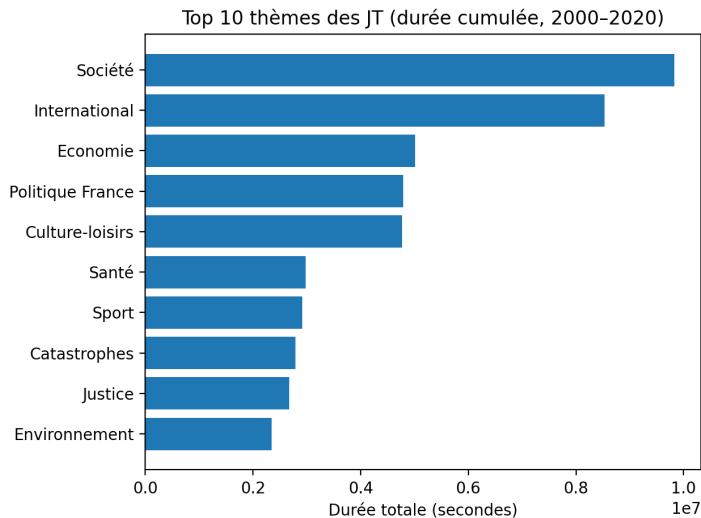


FIGURE 5 – Régression linéaire du nombre d'évocation des partis de gauche et d'extrême droite sur les chaines d'info en continu

2.2 Les JT des chaines généralistes

Top 10 des thèmes des JT



Pour cette première figure, nous avons cherché à identifier les thèmes qui occupent le plus de place dans les journaux télévisés sur le long terme. Pour cela, nous avons regroupé les données par thème et additionné la durée totale associée à chacun d'eux sur l'ensemble de la période 2000–2020.

Nous avons choisi d'utiliser la durée cumulée plutôt que le nombre de sujets, car cet indicateur reflète mieux le temps d'antenne réellement consacré à chaque thème. En effet, un thème peut être abordé très souvent mais de manière brève, alors qu'un autre peut être traité moins fréquemment mais faire l'objet de reportages plus longs.

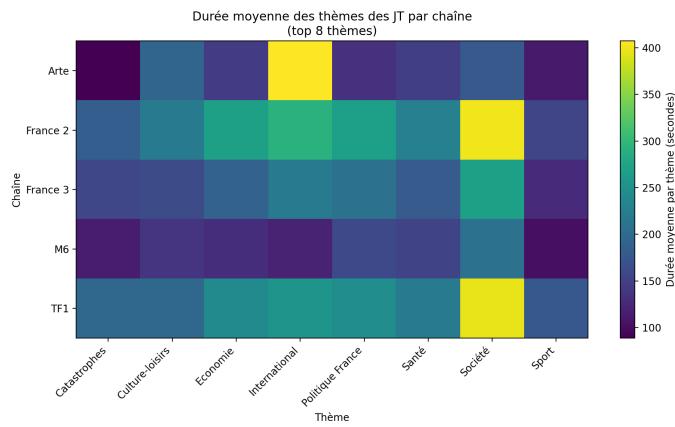
Le graphique montre que le thème Société domine très largement l'ensemble des autres thèmes. Cela suggère que les journaux télévisés accordent une place centrale aux sujets liés à la vie quotidienne, aux faits sociaux et aux enjeux touchant directement la population. Le thème International arrive en deuxième position, ce qui montre l'importance accordée aux événements étrangers, notamment lors de crises, de conflits ou d'événements majeurs.

Les thèmes Économie, Politique France et Culture-loisirs occupent une position intermédiaire. Ils sont présents de manière régulière, mais disposent d'un temps d'antenne plus limité que les thèmes dominants. Cela peut refléter un traitement plus ponctuel ou plus segmenté de ces sujets dans les journaux télévisés.

À l'inverse, des thèmes comme Environnement, Justice ou Catastrophes apparaissent plus bas dans le classement. Leur présence plus réduite en durée cumulée ne signifie pas qu'ils sont absents des JT, mais qu'ils sont souvent traités de manière plus brève ou plus dépendante de l'actualité immédiate. Par exemple, l'environnement peut être davantage couvert lors de certains événements spécifiques, sans constituer un thème structurant permanent.

Dans l'ensemble, cette figure met en évidence une forte concentration du temps d'antenne autour de quelques thèmes principaux. Elle montre que les journaux télévisés reposent sur une hiérarchisation claire de l'information, où certains sujets occupent une place centrale et stable, tandis que d'autres apparaissent de façon plus ponctuelle. Cette structure constitue un point de départ essentiel pour analyser ensuite l'évolution des thèmes dans le temps et les différences de traitement entre chaînes.

Durée moyenne des thèmes par chaîne



Dans un second temps, nous avons cherché à comparer la manière dont les différentes chaînes traitent les principaux thèmes des journaux télévisés. Pour cela, nous avons sélectionné les huit thèmes les plus importants et calculé, pour chaque chaîne, la durée moyenne consacrée à chacun d'eux. Ce choix permet de comparer les chaînes indépendamment du nombre total de journaux diffusés.

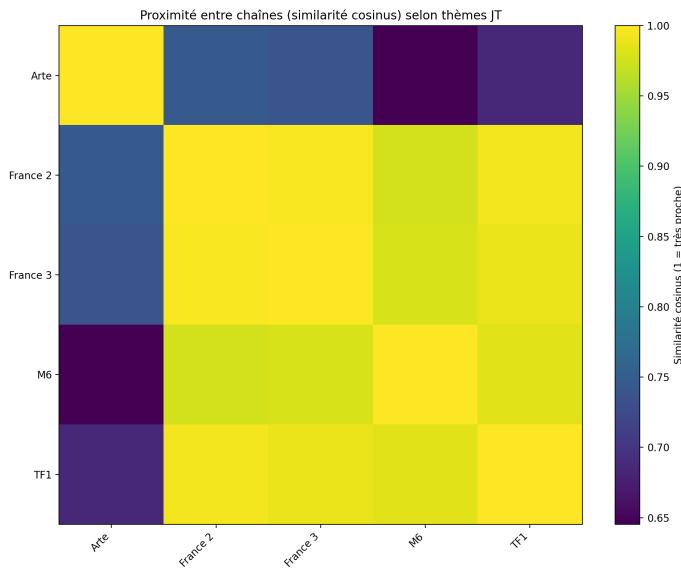
Les résultats sont présentés sous forme de heatmap, car ce type de graphique permet d'identifier rapidement les différences de traitement entre chaînes et entre thèmes, grâce aux variations de couleur.

L'analyse de la figure montre que les chaînes généralistes, en particulier TF1 et France 2, consacrent en moyenne davantage de temps au thème Société. Cela confirme le rôle central de ce thème dans les journaux télévisés, mais aussi une approche assez similaire de l'information entre ces chaînes, qui privilégient des sujets proches du quotidien et des préoccupations du grand public. Le thème International est également très présent sur ces chaînes, même si l'intensité du traitement varie.

À l'inverse, Arte présente un profil différent, avec des durées moyennes plus faibles sur plusieurs thèmes dominants. Cela peut s'expliquer par un format de journal télévisé plus court ou par une approche éditoriale plus ciblée, qui ne repose pas sur une couverture large et répétée des mêmes sujets. M6 se distingue également par des durées moyennes plus faibles, ce qui suggère une structuration différente de ses journaux, possiblement plus rythmée ou plus segmentée.

Cette figure montre ainsi que, même si les chaînes abordent globalement les mêmes grands thèmes, elles ne leur accordent pas le même poids. La différence ne se situe donc pas uniquement dans les sujets traités, mais dans le temps qui leur est consacré. Ces écarts traduisent des choix éditoriaux propres à chaque chaîne, qui influencent la manière dont l'information est hiérarchisée et présentée au public.

Similarité entre chaînes selon les thèmes des JT



L'objectif de cette figure est de comparer les chaînes entre elles afin de déterminer si leurs journaux télévisés présentent des profils thématiques similaires ou, au contraire, des différences marquées. Ici, nous ne cherchons pas à comparer les volumes totaux de diffusion, mais bien la répartition relative des thèmes, c'est-à-dire la manière dont chaque chaîne structure son information.

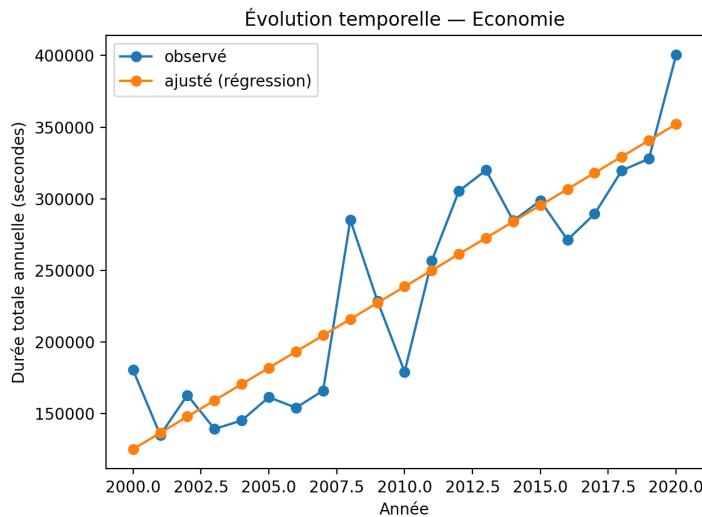
Pour cela, nous avons d'abord regroupé les données par chaîne et par thème, puis additionné la durée totale associée à chaque thème. Cette étape permet d'obtenir, pour chaque chaîne, une vision globale du temps consacré aux différents types de sujets. Afin de rendre les chaînes comparables entre elles, ces durées ont ensuite été transformées en proportions, de sorte que la somme des thèmes représente 100

À partir de ces profils thématiques, nous avons calculé une similarité cosinus entre chaque paire de chaînes. Une valeur proche de 1 signifie que deux chaînes répartissent leur temps d'antenne entre les thèmes de manière très proche, tandis qu'une valeur plus faible indique des choix éditoriaux plus différenciés. Les résultats sont représentés sous forme de heatmap afin de visualiser rapidement les proximités et les écarts entre chaînes.

L'analyse de la figure montre que plusieurs chaînes généralistes, notamment TF1, France 2, France 3 et M6, présentent des profils très proches. Cela signifie que, au-delà des différences de ton ou de mise en forme, leurs journaux télévisés reposent sur une structure thématique très comparable. Ces chaînes traitent les mêmes grands thèmes, dans des proportions similaires, ce qui suggère une certaine standardisation de l'information télévisée généraliste.

À l'inverse, Arte apparaît nettement moins proche des autres chaînes. Cette différence indique un traitement thématique plus spécifique, avec une répartition du temps d'antenne qui s'écarte du modèle dominant des chaînes généralistes. Cela peut refléter une ligne éditoriale plus ciblée ou un positionnement différent vis-à-vis de l'actualité, qui privilégie certains thèmes ou modes de traitement.

Dans l'ensemble, cette figure montre que si les journaux télévisés partagent un socle commun de grands sujets d'actualité, toutes les chaînes ne les hiérarchisent pas de la même manière. Les différences observées ne portent donc pas tant sur les thèmes abordés que sur leur importance relative, ce qui traduit des choix éditoriaux distincts et participe à la diversité – plus ou moins marquée – de l'offre d'information audiovisuelle.



Régression linéaire temporelle du thème Économie

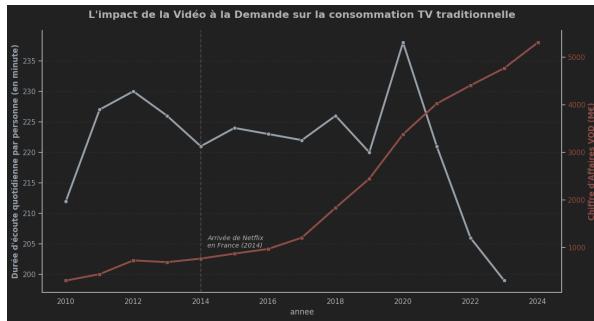
La régression linéaire estimée pour le thème Économie met en évidence une tendance clairement positive et statistiquement significative. Le coefficient associé à l'année est positif et élevé ($+11\,340$), ce qui indique que la durée totale annuelle consacrée à ce thème augmente fortement au fil du temps. La p-value associée à ce coefficient est très faible (inférieure à 1). Le coefficient de détermination R^2 est élevé (environ 0,80). Cela signifie que l'année explique près de 80 % de la variance de la durée totale annuelle.

Interprétation et lien avec les graphiques Ces résultats sont cohérents avec les graphiques d'évolution présentés précédemment, qui montrent une augmentation progressive et assez régulière de la place du thème Économie dans les journaux télévisés. La régression confirme que cette hausse n'est pas seulement visuelle, mais correspond à une tendance structurelle forte sur l'ensemble de la période étudiée. Contrairement au thème Société, dont l'évolution apparaît plus irrégulière et fortement dépendante de l'actualité, le thème Économie semble suivre une trajectoire plus stable dans le temps. Cela suggère une montée en importance durable des sujets économiques dans l'information télévisée, probablement liée à des transformations de long terme (crises économiques, chômage, pouvoir d'achat, mondialisation). La place du thème Économie dans le top 10 des thèmes, combinée à cette tendance temporelle marquée, montre qu'il ne s'agit pas seulement d'un thème important, mais d'un thème dont le poids éditorial s'est renforcé au fil des années dans les journaux télévisés.

3 Evolution du rapport des français à l'information -

3.1 Le constat du délaissement de la télévision

En 2014, l'arrivée de Netflix en France remplace les services de vidéo à la demande historique français comme Orange, Canal+, France TV et TF1. On considère cette date comme l'explosion du marché de la VOD en France. En l'espace de seulement quatre mois, Netflix capture 25,8% des consommateurs français de VOD. Sur le plan financier, la croissance est encore plus frappante. Le marché de la SVOD passe de 29,2 millions d'euros en 2014 à 82,2 millions d'euros en 2015, soit une croissance de +182%. Après ce dernier choc, le marché devient attractif pour d'autres grands acteurs mondiaux comme Amazon Prime en 2016, Apple TV en 2019 ou Disney+ en 2020. Le phénomène est d'autant plus amplifié par le COVID-19 ce qui ancre ces services dans le paysage audiovisuel français. Ceux qui ne voient pas cela d'un bon œil sont les dirigeants des grands groupes historiques de la télévision française qui se sentent menacés et perdent quotidiennement de l'audience. Pour visualiser cela nous allons comparer l'évolution au cours des années la durée d'écoute quotidienne par personne et en minute et l'évolution du chiffre d'affaires de la VOD en millions d'euros.



On constate encore des pics d'augmentation en 2018 (changement dans la récolte des données Médiamétrie) et 2020 (COVID-19) mais globalement la tendance est à la baisse tandis que le chiffre d'affaires de la vidéo à la demande ne cesse d'augmenter. On effectue alors une régression linéaire qui confirme statistiquement cette substitution : le modèle explique 53,2 % de la variance du déclin TV. La variable VOD apparaît comme le moteur principal de cette délinéarisation ($p=0,026$), chaque hausse de revenus du streaming corrélant avec une baisse de l'écoute linéaire. À l'inverse, le Replay n'est pas un prédicteur significatif ($p=0,189$), suggérant que la perte d'audience profite davantage aux plateformes externes qu'au ratrappage interne.

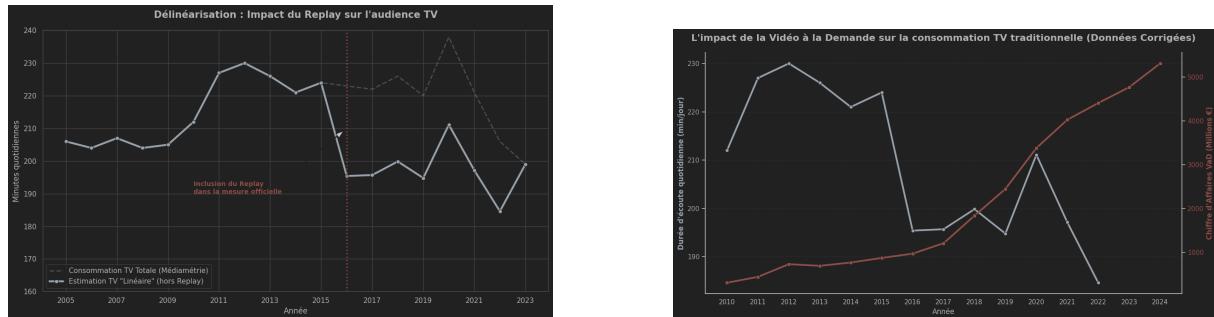
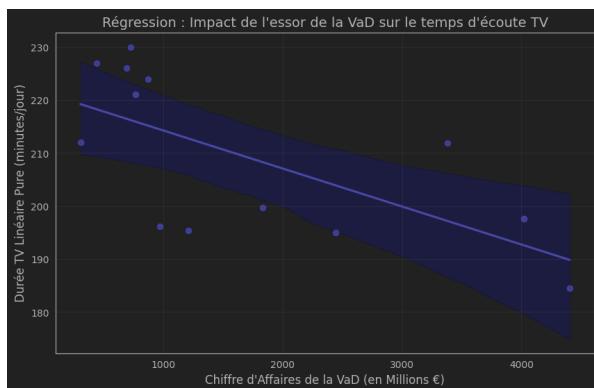


FIGURE 6 – Durée d'écoute TV et Chiffre d'affaires VOD

La façon dont les données ont été récoltées de la part de médiamétrie a évolué au cours des années, pour petit à petit inclure la VOD dans ces données. C'est pour cela que l'on constate certaine augmentation, la plus marquante est en 2016 où toute les VOD sont désormais pris en compte. Pour autant, on constate quand même une grande diminution de la durée d'écoute quotidienne. On va alors raffiné les données en prenant en compte la part de VOD à partir de 2016. À l'inverse, le Replay n'est pas un prédicteur significatif ($p=0,189$), suggérant que la perte d'audience profite davantage aux plateformes externes qu'au ratrappage interne.

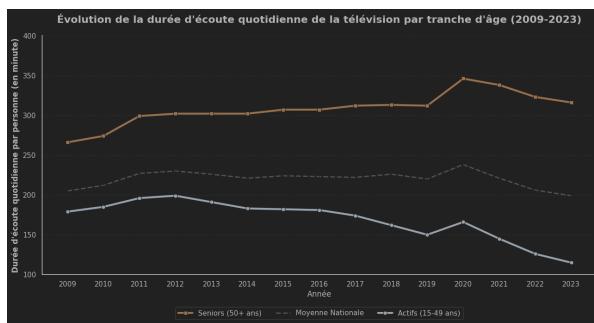


Ce graphique illustre le déclin du flux linéaire et l'essor des plateformes. Une fois isolées de l'effet Replay les données révèlent une chute plus marquée de l'audience traditionnelle. On

observe que les rebonds de 2018 et 2020 ne parviennent pas à inverser la tendance de fond. Cette divergence visuelle confirme que le temps d'écran migre désormais hors du contrôle des diffuseurs historiques au profit d'une consommation de vidéo à la demande. Si l'essor de la VOD explique structurellement la baisse globale de l'audience, cette diminution ne frappe pas toutes les générations avec la même intensité, révélant une fracture générationnelle majeure dans les usages.

3.2 L'Analyse Générationnelle

Cependant, cette érosion du média télévisuel ne frappe pas toutes les générations avec la même intensité. Si le déclin est global, il n'est pas uniforme. Les 15-49 ans se tournent vers fictions à la demande, délaisSENT massivement l'écran traditionnel au profit des plateformes. À l'inverse, les 50 ans et plus maintiennent une fidélité quasi intacte au flux linéaire, devenant le dernier rempart de l'audience historique. Cette fuite des publics actifs vers la VOD place les chaînes françaises devant un dilemme : accepter le vieillissement de leur audience ou transformer radicalement leurs contenus pour tenter de reconquérir leur plus jeune audience.



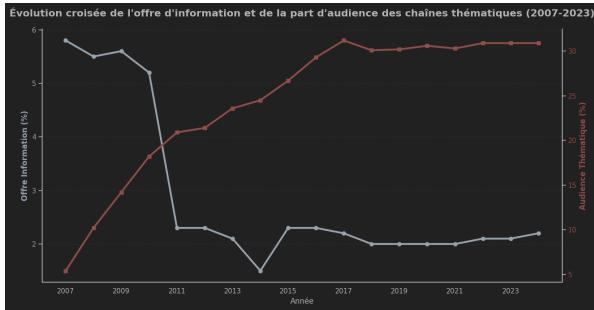
Alors que la courbe des 50+ ans fait preuve d'une résilience remarquable, se maintenant au-delà des 330 minutes quotidiennes, les 15-49 ans amorcent un décrochage sévère dès 2012. Le fossé entre les deux générations explose passant de d'une différence de 120 minutes en 2009 à près de 200 en 2023. Cette dynamique est le point critique pour les diffuseurs historiques. En perdant la tranche des 15-49 ans, la télévision perd sa force de frappe commerciale auprès des annonceurs. Ce "Grand Fossé" prouve que le média ne vieillit pas seulement avec son audience : il perd sa capacité à renouveler ses usagers. La comparaison des modèles de régression entre les tranches d'âge lève les derniers doutes sur la nature de l'évolution en cours. Les chiffres sont sans appel : l'impact de la vidéo à la demande est radicalement opposé selon la génération observée.

Nous avons soumis les données à deux modèles de régression linéaire multiple (OLS), en utilisant le chiffre d'affaires de la VaD et l'offre de fiction comme variables indépendantes. Le modèle affiche un R-squared de 0,888, ce qui signifie que 88,8% de la variation du temps d'antenne des actifs est expliquée par la montée de la VaD et l'évolution de l'offre. Le coefficient de la VaD (-0,0143) agit ici comme le paramètre principal de déclin. Chaque million d'euros supplémentaire injecté dans le marché du streaming réduit la durée d'écoute de cette tranche d'âge proportionnellement à ce coefficient. À l'inverse, le second modèle présente un coefficient positif (+0,0100) pour la VaD. Ce paramètre indique que pour les +50 ans, la vidéo à la demande n'est pas un facteur de diminution mais d'augmentation. La comparaison de ces deux régressions démontre que la VaD est une variable clivante. Pour la cible commerciale (15-49 ans), elle est un prédicteur de rupture, tandis que pour les +50 ans, elle n'est qu'un complément d'usage. Cette divergence de paramètres confirme mathématiquement le vieillissement accéléré de l'audience de la télévision : le média perd ses utilisateurs là où la VaD est forte, et ne les conserve que là où elle n'a pas d'effet de substitution.

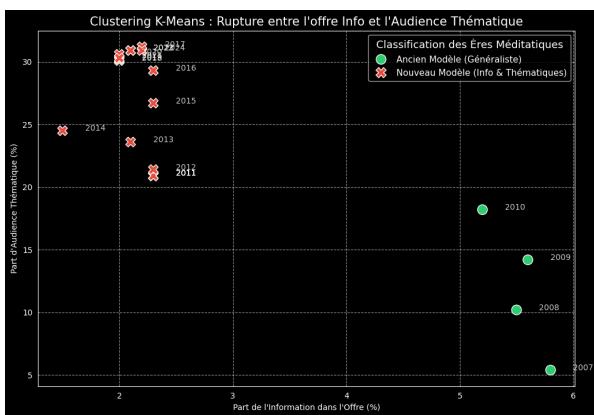
Pour pallier à cette crise les programmateurs de télévision doivent trouver un moyen de se sortir de cette situation en se concentrant sur ce qui ne peut pas être substitué par la vidéo à la demande

3.3 L'information : une Stratégie de Résilience

Face au constat d'un marché de la fiction désormais dominé par les géants du streaming, les groupes audiovisuels français ont dû opérer un pivot stratégique majeur pour assurer leur survie. Il ne reste à la télévision que sa force historique : le direct et le réel. Pour analyser cette mutation, nous avons concentré notre étude sur l'évolution de l'audience thématique (incluant les chaînes d'information en continu comme BFM TV ou CNews) par rapport à l'offre globale de programmes. L'hypothèse est la suivante : la télévision ne cherche plus à divertir par, mais à se spécialiser dans l'actualité. Sur le graphique la proportion de l'offre d'information baisse



puis stagne voire augmente légèrement à partir de 2018. On pourrait croire que cela veut dire que l'offre de journaux télévisés diminue mais si l'on prend en compte que l'offre de télévisions ne cesse d'augmenter, cela veut dire que les programmeurs TV proposent de plus en plus d'information. Et cela est lié à l'augmentation des chaînes thématiques souvent porté uniquement sur l'information (BFMTV, CNews, LCI, France Info). On peut remarqué par ailleurs que même si les audiences télés baissent la proportion d'audience thématique augmente. Cela est révélateur d'une nouvelle stratégie : miser sur les chaînes spécialisées non substituables par les plateformes de vidéo à la demande. Pour valider cette transition d'un modèle à l'autre, nous ne nous sommes pas contentés d'une observation chronologique classique. Nous avons mobilisé un algorithme de Machine Learning, le K-Means Clustering, afin de laisser la donnée identifier elle-même les ruptures structurelles du marché. En croisant la part de l'information dans l'offre et la part d'audience thématique, l'algorithme sépare nos données en deux ères distinctes : une ère "Traditionnelle" à l'offre généraliste et une ère "Spécialisée" où l'information devient le principal levier de croissance et de résilience face au numérique. Ainsi, l'analyse croisée de la consomma-



tion audiovisuelle française entre 2009 et 2023 ne montre pas que la télévision française est en train de s'éteindre. On a prouvé que la vidéo à la demande vole l'audience de la télévision. Un phénomène qui s'accentue chez les 15-49 ans qui préfèrent les plateformes de VOD alors que les 50+ reste sur une demande plus classique. Enfin pour tenter de survivre, le marché de la télévision mise sur ses avantages non substituables, c'est-à-dire une spécialisation dans les chaînes thématiques et plus particulièrement les chaînes d'informations.

4 La représentation des femmes dans les médias -

4.1 L'évolution de la représentation des femmes (TV radio)

Sur la période observée, on voit une progression globale de la part de parole féminine, à la fois à la télévision (2010–2019). Les courbes d'évolution annuelle montrent une hausse plus nette en fin de période, avec une dynamique particulièrement visible à partir du milieu des années 2010, surtout pour certains segments (public/privé).

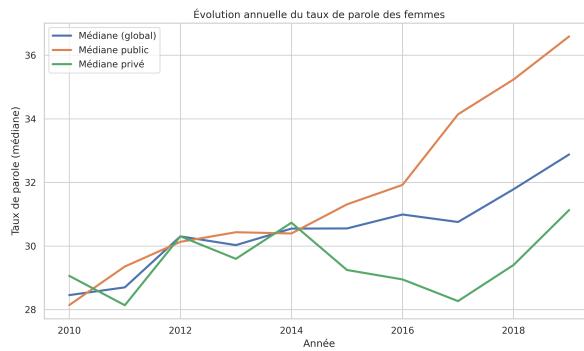
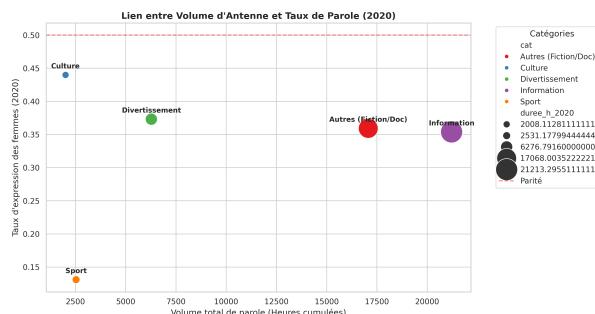


FIGURE 7 – Évolution de la part de parole féminine à la télévision (gauche) et à la radio (droite)

Les régressions confirment statistiquement cette tendance : TV : la part de parole féminine augmente avec l'année (coef. année +0,231; p<0,001), avec un R² très élevé (0,942), ce qui indique que l'année + l'effet chaîne expliquent déjà l'essentiel des variations. la parité ne progresse pas “au hasard”. Elle progresse dans le temps, mais surtout de façon très structurée selon les chaînes/stations (effets fixes massifs). Autrement dit, la représentation dépend fortement de “où” on s’informe, autant que de “quand”.

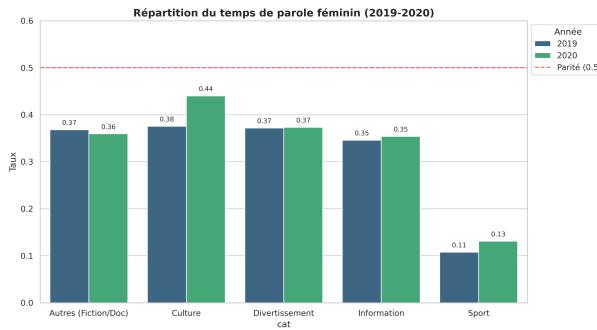


4.2 Impact sur la perception de l'information (diversité / objectivité)

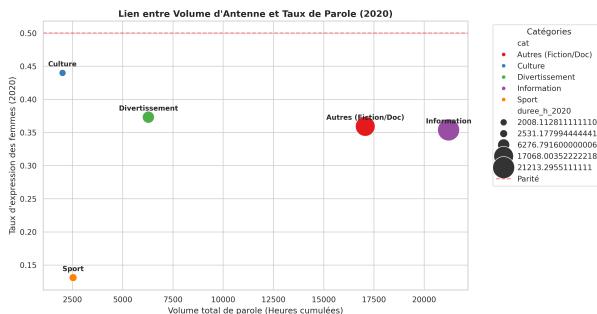
Notre modèle “contextuel” (régression multiple) a un R² faible (0,072) : ça veut dire que l'heure/jour/variables de contexte expliquent une petite partie de la variance totale. Mais même si l'effet global est faible, certains effets sont robustes et significatifs, donc pertinents à discuter : Effet de l'heure : coef. -0,0031 (p<0,001). Concrètement, plus la journée avance, plus la part de parole féminine baisse. Cette tendance est cohérente avec le graphique “par heure”, où les courbes décrochent en fin de journée, période souvent associée à des rendez-vous d'info à forte audience/centralité symbolique. Jours fériés : coef. -0,0098 (p<0,001). Les jours fériés, la visibilité féminine recule légèrement, ce que l'on peut relier à des changements de grilles (programmes spéciaux, sport, rediffusions, etc.), qui modifient la structure des prises de parole.

4.3 Thématiques abordées et lien avec la représentation féminine

Notre comparaison par genres d'émission met en évidence des écarts structurels : par exemple, en 2020, la musique est très au-dessus (0,44), alors que le sport est très bas (0,13), et les journaux/info sont intermédiaires (0,39)



C'est un point clé pour répondre à notre problématique "impact des thèmes diffusés sur la parité" : la parité n'est pas seulement un enjeu de recrutement ou de volontarisme éditorial, c'est aussi un enjeu de répartition des thèmes. Si une chaîne diffuse beaucoup de sport ou de formats historiquement masculinisés, elle peut "tirer" la parité vers le bas, même avec des efforts ponctuels. Autre résultat important : la corrélation entre volume d'antenne et parité est quasi



nulle (corr = 0,05).

5 Conclusion

Ainsi, notre analyse a permis de montrer plusieurs choses. Tout d'abord au niveau thématique on constate une polarisation des thématiques et politique des chaînes, où les sujets régaliens et d'identité sont surreprésentés par rapport aux sujets comme l'environnement et les droits sociaux et une surreprésentation de certains partis, notamment lors de certaines échéances électorales. La télévision n'agit alors pas comme un simple thermomètre qui mesure les opinions des français.es, mais comme un acteur contribuant à l'opinion au travers des thématiques abordées.

Ensuite la télévision traditionnelle fait face à une fracture générationnelle sans précédent. Tandis que les moins de 50 ans délaisse massivement l'écran linéaire au profit de la VOD (Netflix, Prime Video), les plus de 50 ans restent fidèles au flux historique. Pour survivre, les diffuseurs ont opéré un pivot stratégique vers l'information et le direct, des contenus "non substituables" par les plateformes de streaming.

Enfin, même si la part de parole des femmes progresse de manière constante depuis 2010, elle reste très dépendante des thématiques abordées. Certains domaines comme le sport freinent encore la parité globale, et l'on observe une baisse de la visibilité féminine lors des tranches horaires de forte audience en fin de journée.

6 Annexes

- ◆ Code de récupération des données INA : <https://github.com/MaxbanCh/DataScienceProject>
- ◆ Site web du projet : <https://datascience.axithem.fr>
- ◆ Dossier des graphiques utilisés : <https://nextcloud.axithem.fr/index.php/s/agnesawFaD7BXT7>