

Analyse de la couverture médiatique de la campagne présidentielle 2017

Amina Bouchafaa - Maxime Godin - Félix Martel

07/12/2017

Présentation du projet

Au cours de ce projet, nous nous sommes intéressés à la couverture médiatique de la campagne présidentielle française de 2017. Nous avons collecté les articles de trois titres nationaux, L'Express, Le Monde et Le Figaro sur la période du 1er décembre 2016 jusqu'au 31 juillet 2017 à l'aide d'outils de webscraping. Parallèlement, nous avons établi une liste de mots d'intérêt dont nous avons suivi l'utilisation dans les articles de presse retenus. Nous mettons en oeuvre différentes techniques d'analyse et de représentation (nuage de mots, graphiques, analyse factorielle et classification hiérarchique) en les appliquant à des données textuelles mis sous forme de table de contingence.

Création des données

Dans cette partie, nous nous intéressons à la récupération des données à partir des sites Internet des trois titres de presse retenus dans notre étude : Le Monde, L'Express et Le Figaro.

Récupération des articles

La collecte de données à partir de pages Internet est appelée webscraping. Cette technique consiste à exploiter la structure html des pages de sites Internet pour récupérer des données. Nous avons utilisé le package rvest qui permet de récupérer les éléments d'une page web en utilisant des sélecteurs CSS.

Dans un premier temps, nous avons déterminé une liste de mots-clés pour cibler les articles traitant de la campagne présidentielle.

## [1]	2017	affaire	arthaud	asselineau
## [5]	assemblee	assemblée	assistants	bayrou
## [9]	campagne	candidat	candidats	candidature
## [13]	cazeneuve	cheminade	debat	débat
## [17]	denonce	dénonce	deputes	députés
## [21]	droite	dupontaignan	election	élection
## [25]	enquete	enquête	examen	ferrand
## [29]	fillon	gauche	gouvernement	hamon
## [33]	hollande	juge	juppe	juppé
## [37]	justice	lassalle	legislatives	législatives
## [41]	macron	marche	melenchon	mélenchon
## [45]	ministre	montebourg	parlementaires	peillon
## [49]	pen	penelope	penelopegate	politique
## [53]	poutou	president	président	présidentielle
## [57]	présidentielle	primaire	proces	procès
## [61]	programme	sarkozy	sondage	soutien
## [65]	tour	travail	valls	vote
## [69]	voter			

Construction des tables de contingence

Nous avons par ailleurs déterminé une liste de mots d'intérêt dont nous souhaitons mesurer l'utilisation dans les articles retenus.

##	[1]	françois	présidentielle	candidat	macron
##	[5]	fillon	ministre	france	emmanuel
##	[9]	président	tour	primaire	2017
##	[13]	élection	gauche	campagne	droite
##	[17]	ancien	candidats	pen	politique
##	[21]	parti	marine	enquête	hollande
##	[25]	hamon	valls	gouvernement	manuel
##	[29]	travail	benoît	national	marche
##	[33]	république	socialiste	nouveau	affaire
##	[37]	législatives	chef	débat	front
##	[41]	etat	programme	mélenchon	justice
##	[45]	jeanluc	projet	candidate	vote
##	[49]	examen	candidature	nationale	assemblée
##	[53]	républicains	loi	députés	soutien
##	[57]	sondage	emplois	député	épouse
##	[61]	annonce	présidente	victoire	américain
##	[65]	insoumise	philippe	presse	emploi
##	[69]	majorité	penelope	bernard	nicolas
##	[73]	pouvoir	parquet	maire	elysée
##	[77]	électeurs	fictifs	affaires	arnaud
##	[81]	secrétaire	judiciaire	politiques	conseil
##	[85]	mouvement	accord	meeting	état
##	[89]	juge	quinquennat	montebourg	parlementaire
##	[93]	élu	mandat	appel	discours
##	[97]	élections	europe	femme	parlementaires
##	[101]	procès	décision	cour	déplacement
##	[105]	voix	sécurité	parlement	cazeneuve
##	[109]	réforme	vincent	élus	sondages
##	[113]	soupçons	jeune	mesures	millions
##	[117]	droit	retour	voter	vainqueur
##	[121]	afp	européen	attentat	dénoncé
##	[125]	famille	ligne	européenne	propositions
##	[129]	canard	visant	penelopegate	résultats
##	[133]	défense	scrutin	bayrou	société
##	[137]	union	sarkozy	enchaîné	bfmtv
##	[141]	tf1	accusé	police	public
##	[145]	rapport	peillon	intérieur	modem
##	[149]	camp	peine	leader	économie
##	[153]	sociaux	jugé	révélations	prison
##	[157]	critiques	circonscription	opposition	economie
##	[161]	risque	condamné	financier	femmes
##	[165]	droits	fonds	fondateur	voté
##	[169]	crise	autorité	lutte	syndicats
##	[173]	policiers	électorale	présidence	ministère
##	[177]	tribunal	centre	extrême	jeunes
##	[181]	passé	économique	confiance	alliance
##	[185]	porteparole	bruno	socialistes	présumés
##	[189]	bilan	chômage	social	prix
##	[193]	ministres	avocat	plainte	critiqué

## [197]	ferrand	accusations	soupçonné	rassemblement
## [201]	richard	participation	populaire	violences
## [205]	histoire	écologiste	conseiller	bordeaux
## [209]	juppé	laurent	fictif	attaque
## [213]	pierre	intentions	retrait	présumé
## [217]	sénat	menace	surprise	journalistes
## [221]	militants	investiture	matignon	cérémonie
## [225]	interrogé	soutenir	juges	défendre
## [229]	défendu	victimes	défaite	épreuve
## [233]	scandale	briguer	guerre	entredoux
## [237]	christophe	instruction	soutiens	comptes
## [241]	édouard	polémique	éllysée	investi
## [245]	jacques	terrorisme	proposition	khomri
## [249]	promis	assistants	système	préliminaire
## [253]	démission	enquêteurs	corruption	expremier
## [257]	michel	budget	entourage	hommage
## [261]	favori	ordre	lrem	terroriste
## [265]	cgt	tribune	parrainages	futur
## [269]	présidentiel	banque	débats	patrick
## [273]	avenir	commission	stratégie	procédure
## [277]	renoncement	responsables	viol	dénonce
## [281]	attentats	action	rtl	attaques
## [285]	couple	charge	thomas	finalistes
## [289]	argent	meurtre	intention	exministre
## [293]	députée	yannick	fraude	visé
## [297]	fiscale	jeanmarie	jdd	interpellation
## [301]	révélé	dupontaignan	régime	jeanchristophe
## [305]	forces	policier	bataille	blanc
## [309]	christian	placé	programmes	brigitte
## [313]	retraite	udi	financement	doute
## [317]	faux	bruxelles	échec	école
## [321]	recours	figaro	adversaire	aulnaysousbois
## [325]	morts	direction	avocats	fort
## [329]	education	milliards	faute	absence
## [333]	gérard	sortant	croissance	étranger
## [337]	seinesaintdenis	éliminé	européens	libération
## [341]	convaincre	protection	jadot	américains
## [345]	sénateurs	négociations	électoral	centriste
## [349]	étrangères	frais	rassembler	rue
## [353]	allemand	liberté	parole	poutine
## [357]	résultat	ferme	assistante	nathalie
## [361]	actions	debout	cfdt	congrès
## [365]	battu	sceaux	champion	viceprésident
## [369]	référendum	déclarations	dame	angela
## [373]	merkel	transparence	ralliement	sénateur
## [377]	angeles	islamique	internauts	sud
## [381]	amf	harris	chiffres	accuse
## [385]	éducation	allemande	anciens	air
## [389]	avis	radio	dialogue	mère
## [393]	économiques	dirigeants	anne	marion
## [397]	marc	sociales	vigueur	vision
## [401]	clients	prétendants	exécutif	fédération
## [405]	russes	revenus	finances	cambadélis
## [409]	abus	fonction	provoqué	aide

## [413]	partisans	procureur	odoxa	territoires
## [417]	équipes	soupçonnés	mariage	partis
## [421]	évoque	antiterroriste	jeux	philippot
## [425]	documents	ami	abstention	statut
## [429]	turquie	défend	concurrents	logement
## [433]	thierry	yvelines	promesses	laïcité
## [437]	dynamique	promet	renoncé	barack
## [441]	locaux	réformer	colère	parcours
## [445]	concours	jeanfrançois	zone	accueil
## [449]	maladie	honneur	bercy	mari
## [453]	royaumeuni	usine	empêtré	chirac
## [457]	solidarité	réussi	phare	analyse
## [461]	essentiel	immobilier	oeuvre	australie
## [465]	compétition	combat	intérêt	succès
## [469]	arrondissement	migrants	filles	mitterrand
## [473]	frontiste	ifop	popularité	projets
## [477]	mission	nomination	centrale	benoit
## [481]	florian	médiatique	fils	ligue
## [485]	espoir	myriam	judiciaires	lambert
## [489]	fustigé	militaires	jean	rémunération
## [493]	engagé	critique	maintien	rival
## [497]	soutenu	agression	républicain	changement
## [501]	alliés	commune	dénonçant	portrait
## [505]	autorités	citoyens	sens	parents
## [509]	audience	réagi	engagement	arme
## [513]	italie	sanctions	blesse	électorales
## [517]	culture	suisse	primaires	opposants
## [521]	rejeté	communication	avance	contrat
## [525]	rmc	adversaires	domicile	mur
## [529]	versailles	lice	soupçonnée	climat
## [533]	maréchal	mathieu	collaborateurs	suppléant
## [537]	mondes	élue	voile	villes
## [541]	lagarde	investigations	radical	mandats
## [545]	réalité	travailler	problème	qualification
## [549]	impôt	erreur	perdre	responsable
## [553]	sursis	détention	affiche	constitutionnel
## [557]	reportage	influence	arrêt	csa
## [561]	cabinet			

Analyses et résultats

Utilisation de mots d'intérêt dans la presse

Nuages de mots

Nous avons utilisé des nuages de mots pour représenter le vocabulaire utilisé dans les trois quotidiens au cours de chaque mois de la période retenue.

Mois	Le Monde	Le Figaro	L'Express
------	----------	-----------	-----------



déc.

Evolution temporelle lissée

Analyse factorielle des correspondances

Représentation graphique des articles et des mots d'intérêt

Classification hiérarchique sur composantes principales