

1 sur 3

Dans les abréviations, on retrouve également : “mod” pour “modéré”, notamment dans l’expression “prix modéré” ; “ecr” pour “écrire à” ; “pet.” pour petite maison, petit appartement et petit loyer ; “gar.” pour “garni”, comme “chambre garnie” ; “com.” notamment pour “comme”.

Au niveau de l’analyse des sentiments, on retrouve principalement des phrases subjectives, positives ou négatives, pour les phrases complètes, et des phrases neutres et objectives pour les petites annonces.

Phrases	Négatif	Positif	Subjectif	Objectif
Avant de faire démarrer le camion sur lequel le coffre-fort avait été hissé, on tendit sur deux perches une grande pancarte avec cette inscription : «Le coffre-fort aux cent millions».		30%	20%	
C'est un avocat, M. Emilio Cotarello y Mori, membre de l'Académie royale d'Espagne, qui habitait à Madrid, 33, rue Ferraz, l'étage au-dessus de celui qu'occupaient les Humbert.		20%	10%	
Un jour où j'allais rendre visite à une dame d'Auteuil qui s'occupait beaucoup de l'Œuvre de la Première Communion — Mme Fanny Dubois-Davesnes — je lui demandai comment était mort l'abbé Roussel		30%	15%	
Allez dormir et 11e rêvez pas à ceux que vous connaissez, amis ou ennemis.		8%	62%	
Pour les hommes, peu importe, leur enveloppe toujours quelconque n'a nulle influence extérieure, tandis que pour nous tout est là, et savez-vous pourquoi l'année qui s'éteint a été plutôt néfaste et pourquoi plusieurs de celles qui l'ont précédée ont été si peu reluisantes?	9%		24%	
Ceintures pv dames, modèles divers, ; au lieu de 1.85 'Taffetas noir, très belle qualité, sans concurrence.		26%	24%	
L'auteur de l'Oreille coupée, parlant devant un auditoire de jeunes filles, et lui expliquant les beautés du théâtre d'Augier, quelle occasion de scabreuses allusions, de piquantes anecdotes !		20%	20%	
Une bonne maison de rentier A-Laokon, rue du Champ-de-l'Eglise; 97; suivie de jardin avpc-volières, contenant on sùpèrficio 2 ails 99 centiares		70%	70%	
149 LUSTRES D'OCCAS.®: Achat et vente à la fabrique dolus:res.l29, rue Brogniez, près rue Tête de Mouton9742	0%	0%		100%
Persuadé qu'il ne se : relèverait pas; .il fit prévenir M 110 Lemerrier, qui était sa seule héritière.	10%		20%	

L'apport du topic modeling

D'après la présentation de l'outil gensim, le topic modeling permet de réaliser une représentation vectorielle de l'analyse sémantique d'un corpus.

Dans le cadre de l'exploration du corpus CAMille, une telle représentation permettrait d'identifier la répétition de structures de phrases, tant du côté des phrases complètes, que du côté des nombreuses petites annonces.