

Décrire l'offre commerciale comme espace social

Une enquête exploratoire sur neuf espaces commerciaux franciliens

Master Sciences sociales - Parcours Quantifier en sciences sociales

2025-2026

MÉMOIRE DE RECHERCHE

Soutenu par

Zhuoran LI

Session

Juin 2026

Directeur

Jérôme Deauvieu

Rapporteurs

Étienne Penissat, Jérôme Deauvieu, Nathan Lévêque

Table des matières

Introduction	1
Chapitre 1. Du commerce urbain aux variables sociales : ce que les outils statistiques voient et ce qui leur échappe	6
1.1. Le centre commercial comme espace de classement social	6
1.2. Les outils statistiques actuels du commerce urbain	9
1.3. La NAF : convention statistique et ses limites pour l'analyse sociale	11
1.4. Construire de nouvelles variables : enjeux, hypothèses et limites	12
Chapitre 2. Construction et exploration des données	15
2.1. Le terrain : neuf espaces commerciaux franciliens	15
2.2. Sources et collecte	17
2.3. Le codage du discours : trois schèmes	18
2.4. Variables analytiques	21
2.5. Limites de la production des données	23
Chapitre 3. L'enrichissement à l'épreuve des prix : typologie, grammaire et sélection spatiale	24
3.1. Vue d'ensemble du corpus	24
3.2. L'espace factoriel des registres discursifs	27
3.3. Le discours et la hiérarchie des prix	32
3.4. Amplitude interne de la gamme observée : plancher, plafond et extension de gamme	37
3.5. Position dans le secteur et profils de prix synthétiques	39
3.6. Marques transversales et position locale	41
3.7. La sélection spatiale et l'hétérogénéité des grands magasins	42
3.8. Ce que la base permet de retenir	44
Conclusion	45
Bibliographie	47
Références scientifiques	47

Autres références	49
Sources et bases de données.....	49
Annexes	50
Annexe A. Exemples de codage du discours	50
Annexe B. Figures supplémentaires de l'analyse des correspondances multiples.....	51
Annexe C. Modèles linéaires de probabilité (M7 et M8).....	54
Annexe D. Distribution détaillée du corpus (type d'espace, secteur, cohorte de fondation)	55
Annexe E. Tableau récapitulatif des variables construites.....	56

Introduction

Un centre commercial ne montre pas simplement la ville autour de lui. Il sélectionne certaines enseignes, certains niveaux de prix et certains voisinages marchands. Et tous ces espaces ne montrent pas la même ville. Au Bon Marché Rive Gauche, on trouve Hermès, Chanel, Saint Laurent, Loewe, des marques de niche dans la parfumerie et l'habillement. À Westfield Vélizy 2, dans les Yvelines, l'offre se concentre sur des enseignes de milieu et haut de gamme comme Sandro, Maje, COS, Calvin Klein ou Tommy Hilfiger. À Westfield Rosny 2, en Seine-Saint-Denis, l'offre fait davantage apparaître des enseignes comme Kiabi, Géo, La Halle ou Jennyfer. Ces différences sont assez immédiates pour quiconque fréquente ces lieux. Mais certaines marques traversent aussi des univers commerciaux a priori distincts. La base constituée pour ce mémoire montre par exemple que Levi's est présente dans huit des neuf centres étudiés, et qu'Adidas apparaît dans six centres, allant du Bon Marché à Aéroville.

C'est en regardant ces marques présentes dans plusieurs centres que mon problème de recherche s'est précisé. Levi's au Bon Marché et Levi's à Rosny 2 renvoient à une même marque et à une gamme de produits largement comparable. Pourtant, l'enseigne ne joue pas le même rôle dans les deux assortiments. Au Bon Marché, elle fait partie des enseignes les plus accessibles d'un assortiment dominé par le luxe ; dans un centre commercial de couronne, elle peut au contraire figurer parmi les enseignes relativement hautes d'une offre davantage structurée par les enseignes standard ou discount. La même enseigne, à prix relativement comparables, peut donc ne pas occuper la même position relative dans l'espace de l'offre. D'autres enseignes très diffusées, comme Sephora, relèvent d'un cas légèrement différent : elles apparaissent dans plusieurs centres de couronne, sans pour autant traverser les grands magasins centraux du corpus. Elles signalent donc moins une circulation entre univers commerciaux opposés qu'une forte stabilisation dans l'espace des centres commerciaux périphériques.

C'est ce décalage qui m'a conduit à déplacer l'unité d'analyse. Compter les boutiques ne suffit pas ; il faut regarder les voisinages, les écarts de gamme et la position relative des enseignes. Un centre commercial organise en effet des coprésences : une marque y prend sens par les enseignes qui l'entourent, par les prix auxquels elle se compare et par les contrastes qu'elle produit dans l'assortiment. Cette position tient à la fois à un segment de prix, à un type

de discours, à un certain prestige et au voisinage des autres enseignes présentes sur le même site. L'enquête porte donc moins sur les produits pris isolément que sur ces voisinages commerciaux : quelles enseignes sont réunies, quelles gammes coexistent, quels contrastes naissent de leur coprésence, et comment chaque assortiment forme un ensemble cohérent ou contrasté, reconnaissable dans l'espace francilien du commerce.

La cohérence interne de ces assortiments peut être abordée à partir de la notion foucauldienne d'hétérotopie¹ : Foucault désigne par là des lieux réels qui juxtaposent en un seul espace plusieurs emplacements et y agencent des éléments hétérogènes selon une logique propre. Coëffé et Morice la reprennent à propos du grand magasin parisien. Je la mobilise ici de façon limitée : elle me sert à formuler un problème d'agencement. Comment un même lieu réunit-il des enseignes très différentes tout en produisant un assortiment lisible ?²

Cette question fait apparaître un décalage dans les outils disponibles. Les travaux quantitatifs sur le commerce urbain en France comptent les établissements, mesurent leur densité, distinguent centres-villes et périphéries. L'Insee identifie ainsi des pôles commerçants par algorithme de densité et suit leur dynamique territoriale³. Antoine Fleury, Hélène Mathian et Thérèse Saint-Julien⁴, dans une étude sur Paris intra-muros, définissent et classent des centralités commerciales à partir de la concentration et de la spécialisation de l'offre. Ces approches sont pertinentes pour décrire la couverture territoriale et la hiérarchie fonctionnelle du commerce, mais elles restent peu sensibles au positionnement social des enseignes. Le code NAF 47.71Z, qui classe les commerces de détail d'habillement en magasin spécialisé, peut ainsi ranger sous une même catégorie des commerces socialement très éloignés, d'une boutique de luxe à une enseigne populaire. Les enquêtes sociologiques et ethnographiques sur les mondes du commerce et les usages sociaux des galeries marchandes éclairent les pratiques

¹ Foucault Michel, « Des espaces autres », *Empan*, 2004, vol. 54, n° 2, p. 12-19.

² Coëffé Vincent et Morice Jean-René, « The Parisian department store as a paradigmatic place for interactions between tourism and shopping: the production of a heterotopia », *Belgeo. Revue belge de géographie*, 2020, n° 1.

³ Cazaubiel Arthur et Guymarc Gaël, « La déprise du commerce de proximité dans les centres-villes des villes de taille intermédiaire », *Insee Première*, 2019, n° 1782, <https://www.insee.fr/fr/statistiques/4248184>, consulté le 3 juin 2026.

⁴ Fleury Antoine, Mathian Hélène et Saint-Julien Thérèse, « Définir les centralités commerciales au cœur d'une grande métropole : le cas de Paris intra-muros », *Cybergeo*, 2012.

et les sociabilités marchandes, notamment chez Vincent Chabault⁵, mais ne prennent pas directement pour objet, dans la perspective adoptée ici, la composition systématique de l'offre. Ce qui manque surtout, pour mon objet, est une mesure de l'assortiment lui-même : non seulement combien de commerces, mais lesquels, à quels prix et dans quel voisinage. La position sociale relative des enseignes à l'intérieur d'un centre reste rarement objectivée.

Le mémoire prend ce décalage pour point de départ. Par centre commercial, j'entends ici un espace marchand, le plus souvent fermé, qui regroupe plusieurs commerces sous un toit commun et dont la composition est gérée comme une entité unique. Cette définition réunit sous un même concept des objets que les nomenclatures professionnelles séparent : l'Insee désigne par exemple séparément le *grand magasin*, établissement de vente au détail dont la surface de vente dépasse deux mille cinq cents mètres carrés et qui réalise moins du tiers de son chiffre d'affaires en produits alimentaires⁶. Cette distinction administrative est légitime, mais elle laisse au second plan une continuité historique soulignée par la littérature sur le commerce urbain. Walter Benjamin, dans son projet inachevé sur les passages parisiens⁷, a fait des arcades du XIX^e siècle l'un des lieux d'observation privilégiés de la modernité marchande : un espace fermé, vitré, à la fois rue et magasin, où la marchandise se donne à voir dans un dispositif spatial spécifique. On peut inscrire les grands magasins du Second Empire, puis les centres commerciaux périphériques développés à partir des années 1960, dans une même généalogie des espaces commerciaux modernes, sans y voir un développement linéaire. Dans les cas étudiés, l'offre n'est pas une simple addition spontanée de commerces : elle correspond à un assortiment sélectionné et piloté dans un espace enveloppant. C'est cette propriété commune qui rend la comparaison possible entre Le Bon Marché, les centres Westfield et le BHV Parly 2. Les travaux de Jon Goss⁸ sur le mall

⁵ Chabault Vincent, *Sociologie du commerce*, Paris, Armand Colin, 2024, chap. 6, « Le commerce et le territoire », p. 224-291.

⁶ Insee, « Grand magasin », Lexique de la statistique publique, en ligne, <https://www.insee.fr/fr/metadonnees/definition/c1799>, consulté le 28 mai 2026.

⁷ Benjamin Walter, *Paris, capitale du XIX^e siècle. Le livre des passages*, Paris, Éditions du Cerf, 1989, p. 47-59.

⁸ Goss Jon, « The “Magic of the Mall”: An Analysis of Form, Function, and Meaning in the Contemporary Retail Built Environment », *Annals of the Association of American Geographers*, 1993, vol. 83, n° 1, p. 18-47.

américain et de Sharon Zukin⁹ sur les paysages urbains de la consommation permettent de prolonger cette intuition : les espaces commerciaux ne sont pas seulement des contenants, leur forme matérielle participe aussi à l'orientation des pratiques. La distinction entre grand magasin et centre commercial ne disparaît pas pour autant : elle est reprise au chapitre 2 comme l'une des dimensions de la typologie spatiale, sans structurer le découpage de l'objet.

Ce que j'appelle offre commerciale, ce sont les enseignes présentes dans un centre, mais aussi leurs caractéristiques : segment tarifaire, registre discursif, rattachement juridique ou administratif, ancienneté, type d'activité. L'unité d'observation que je retiens est le couple enseigne × centre. Si Levi's est dans huit centres de mon échantillon, il apparaît huit fois dans la base. Ce niveau d'observation tient directement au cas évoqué plus haut : une même enseigne ne joue pas le même rôle selon le centre où elle se trouve. J'ai donc codé l'offre au niveau du couple enseigne × centre, pour pouvoir garder cette différence de position.

La question générale qui sous-tend l'enquête est l'hypothèse formulée par Pierre Bourdieu dans *La Distinction*¹⁰ : il existe une homologie structurale entre la position des agents dans l'espace social et leurs pratiques de consommation, qui se traduit par une correspondance entre la stratification de l'espace social et la différenciation des pratiques, des goûts, des biens et des styles de vie. Si l'on transpose cette intuition à l'offre commerciale, on peut faire l'hypothèse que la composition des centres commerciaux franciliens est organisée de manière à correspondre à la structure sociale des territoires où ils s'implantent. Je travaille donc sur l'offre disponible, et non sur les consommateurs : les enseignes présentes, leurs prix, leurs discours et leur position dans chaque assortiment. J'examine si ces assortiments dessinent eux-mêmes une structure différenciée, susceptible d'entrer en correspondance avec les espaces sociaux où ils s'implantent. Un tel test supposerait un terrain plus large, davantage de variables territoriales et une comparaison avec d'autres espaces commerciaux. Le présent mémoire se limite à une première opérationnalisation empirique de cette hypothèse. Il construit une base enseigne × centre et examine ce que cette base fait apparaître lorsque les

⁹ Zukin Sharon, *Landscapes of Power: From Detroit to Disney World*, Berkeley, University of California Press, 1991, p. 3-23.

¹⁰ Bourdieu Pierre, *La Distinction. Critique sociale du jugement*, Paris, Éditions de Minuit, 1979, chap. 3, « L'habitus et l'espace des styles de vie », p. 189-248.

assortiments sont mis en relation avec les prix observés, les discours de marque et les types d'espace commercial.

La problématique du mémoire peut donc se formuler ainsi : comment les espaces commerciaux franciliens produisent-ils un ordre social de l'offre, en y réunissant des enseignes dotées de prix, de discours de marque et de positions relatives différenciés ? Et à quelles conditions peut-on construire une base quantitative permettant d'objectiver cet ordre, au-delà des catégories administratives ordinaires comme la NAF ? Le travail comporte deux niveaux. Le premier est méthodologique : il porte sur les catégories disponibles, leurs limites, et les choix qu'il faut faire pour construire de nouvelles variables. Le second est empirique : il s'agit d'appliquer ces choix à un échantillon concret et d'en examiner les premiers résultats. Dans l'enquête, les deux niveaux se sont construits ensemble. Dans les faits, l'enquête a progressé par allers-retours : la collecte des listes d'enseignes a fait apparaître les limites de la NAF, tandis que la lecture de Desrosières¹¹ a permis de problématiser ces limites comme un enjeu de catégorisation statistique. La réflexion sur les catégories utilisées ne vient donc pas après la mesure : elle accompagne la constitution même de la base, puisque les variables disponibles ne décrivent qu'imparfaitement le type de différenciation que l'enquête cherche à observer.

Le terrain s'organise autour de neuf espaces commerciaux franciliens : trois grands magasins centraux parisiens (Le Bon Marché Rive Gauche, La Samaritaine, BHV Rivoli), quatre espaces commerciaux de couronne (Westfield Les 4 Temps à La Défense, Westfield Vélizy 2, BHV Parly 2 et Westfield Rosny 2), un centre commercial généraliste (Italie 2 dans le 13^e arrondissement de Paris) et un espace de commerce de transit (Westfield Aéroville, près de Roissy-Charles-de-Gaulle). La base finale rassemble 2 978 observations enseigne × centre, correspondant à 2 173 marques uniques incluses dans la base. Pour chaque marque, le discours institutionnel a été extrait depuis le site officiel, lorsqu'il était disponible, puis codé selon trois schèmes distincts : une forme discursive dominante (luxue patrimonial, création d'auteur, standard), une intensité d'enrichissement (forte, faible, nulle), et douze registres lexicaux codés, dont onze conservés comme registres actifs après exclusion d'un registre trop rare. Les prix ont été collectés à partir des catalogues publics des marques, lorsque ceux-ci

¹¹ Desrosières Alain, *La Politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte, 1993, p. 9-16 et p. 397-418.

étaient accessibles, en conservant jusqu'à dix prix observables par marque afin de construire un prix d'entrée, un prix médian et un prix maximal. La position spatiale a été enregistrée pour chaque observation, et la base SIRENE de l'Insee a été interrogée pour obtenir les informations administratives et juridiques disponibles.

Le plan du mémoire avance en trois temps. Le premier chapitre revient sur les outils statistiques qui décrivent aujourd'hui le commerce urbain en France et examine ce qu'ils permettent et n'autorisent pas de voir. Cette discussion, ancrée dans la sociologie de la quantification, sert à justifier la construction de nouvelles variables. Le second chapitre présente le protocole d'enquête : le périmètre du terrain, la chaîne de constitution de la base et les choix qui ont posé problème. Le troisième chapitre propose une première lecture de la base, qui déploie une analyse factorielle de l'espace des registres discursifs, examine les corrélats économiques de ces registres par régression, et teste si les types d'espace commercial fonctionnent comme dispositifs de sélection des formes discursives. Une conclusion revient sur les régularités empiriques observées et sur les pistes d'extension.

Chapitre 1. Du commerce urbain aux variables sociales : ce que les outils statistiques voient et ce qui leur échappe

1.1. Le centre commercial comme espace de classement social

L'objet de l'enquête est d'abord la position des enseignes dans un assortiment. Une enseigne installée dans un centre n'y apparaît pas comme une unité isolée : sa position dépend des marques qui l'entourent, des écarts de prix avec ses voisines et des contrastes de discours qui la situent dans l'assortiment. Par classement social, j'entends donc moins une opération explicite de tri des consommateurs que la production d'un ordre entre des enseignes, des prix, des styles de discours et des formes de prestige. Décrire systématiquement cette composition permet d'observer la différenciation sociale telle qu'elle s'inscrit dans l'offre commerciale.

Le fait de mettre certaines enseignes ensemble, et pas d'autres, produit déjà une forme de hiérarchie. C'est ce travail de classement, interne à l'offre, que ce chapitre cherche à décrire.¹²

Le problème devient alors celui d'une possible correspondance entre espace de l'offre et espace social, ce que Bourdieu nomme homologie. Transposée au commerce,¹³ cette hypothèse conduit à analyser la composition d'un centre comme le résultat d'ajustements économiques et symboliques au territoire où il s'inscrit, et comme un indice indirect des relations entre offre commerciale et différenciation territoriale.

Cette transposition pose quand même un problème, parce que mon objet n'est pas directement celui de Bourdieu. L'hypothèse de *La Distinction* porte sur la relation entre la position des agents dans l'espace social et leurs pratiques de consommation : elle suppose des sujets, leurs goûts et leurs dispositions. Or mon enquête porte sur l'organisation de l'offre elle-même, et non sur les consommateurs : la manière dont les enseignes se distribuent et se hiérarchisent à l'intérieur d'un espace commercial. C'est donc moins l'homologie des goûts que la logique de champ qui éclaire mon objet : un espace structuré de positions, où chaque enseigne se définit relativement aux autres. C'est dans cette direction que Bourdieu a lui-même prolongé son cadre du côté de l'économie, en analysant le marché comme un champ structuré par la position des producteurs¹⁴. Je retiens alors de l'homologie une intuition structurale : l'idée qu'un espace de biens peut être organisé de manière différenciée et entrer en correspondance avec la différenciation des territoires. La composition d'un centre n'est jamais la projection mécanique d'une structure sociale préexistante ; elle peut correspondre au territoire environnant comme le décaler ou le simplifier. Mon enquête examine empiriquement comment les enseignes se distribuent dans neuf centres concrets, une fois mises en relation avec les prix observés et les types d'espace commercial. Le cas de Levi's, évoqué en introduction, dit assez bien ce que je retiens de l'homologie : la gamme bouge peu d'un centre à l'autre, c'est le voisinage qui change, et donc la position relative. C'est à ce niveau que

¹² Foucault Michel, « Des espaces autres », *Empan*, 2004, vol. 54, n° 2, p. 12-19 ; Coëffé Vincent et Morice Jean-René, « The Parisian department store as a paradigmatic place for interactions between tourism and shopping: the production of a heterotopia », *Belgeo. Revue belge de géographie*, 2020, n° 1.

¹³ Bourdieu Pierre, *La Distinction. Critique sociale du jugement*, Paris, Éditions de Minuit, 1979, p. 196-198, notamment « L'homologie entre les espaces ».

¹⁴ Bourdieu Pierre, *Les structures sociales de l'économie*, Paris, Seuil, 2000, p. 233-270, en particulier « Le champ économique », p. 233-240.

l'intuition bourdieusienne me semble utile : celui d'un espace de positions où chaque enseigne se définit par rapport aux autres. Je reprends l'homologie comme hypothèse de structuration de l'offre. L'analyse porte sur les positions relatives des enseignes dans les assortiments ; elle ne tranche pas la correspondance avec les consommateurs. C'est dans ce sens que je mobilise aussi Roueff, lorsqu'il invite à regarder le travail concret des intermédiaires.¹⁵

Ce travail de classement devient observable dans les propriétés attribuées aux marchandises elles-mêmes. Luc Boltanski et Arnaud Esquerre, dans *Enrichissement*¹⁶, proposent d'analyser les marchandises à partir de formes de valorisation. Le prix d'un bien ne renvoie pas seulement à un coût ou à une utilité : il dépend aussi des opérations par lesquelles l'objet est présenté, qualifié, rattaché à une histoire, à une origine, à une rareté ou à un usage. Les auteurs distinguent quatre formes de valorisation (standard, actif, collection et tendance) qui reposent sur des modes de présentation différents, analytiques ou narratifs, et sur des rapports distincts au temps et à la valeur. La forme *collection* retient particulièrement mon attention, parce qu'elle rattache l'objet à un récit ancré dans le passé et tire sa valeur de la mémoire attachée à l'ensemble dont il fait partie. Dans l'enquête, les discours institutionnels sont donc traités comme des indices du travail de valorisation des marques. Tous ne relèvent pas de l'enrichissement au même degré : certaines organisent entièrement leur récit autour du patrimoine ou de la création, tandis que d'autres n'empruntent que quelques marqueurs isolés à ces registres. Le prix, à lui seul, ne dit pas comment une marque justifie sa position : deux enseignes peuvent vendre cher autour de récits opposés, l'une de l'héritage, l'autre de la création. C'est cette différence que je cherche à coder en m'appuyant sur l'enrichissement, là où une nomenclature d'activité ne distingue que des secteurs.

Le codage du discours pose une question de validité : que mesure-t-on lorsqu'on repère, dans le discours d'une marque, un registre patrimonial ou un registre du care ? La distinction proposée par Michèle Lamont et Virág Molnár¹⁷ entre frontières symboliques (les distinctions

¹⁵ Roueff Olivier, « Les homologies structurales : une magie sociale sans magiciens ? La place des intermédiaires dans la fabrique des valeurs » dans *Trente ans après La Distinction, de Pierre Bourdieu*, Paris, La Découverte, 2013, p. 153-164.

¹⁶ Boltanski Luc et Esquerre Arnaud, *Enrichissement. Une critique de la marchandise*, Paris, Gallimard, 2017, p. 153-197, p. 243-285 et p. 294-314.

¹⁷ Lamont Michèle et Molnár Virág, « The Study of Boundaries in the Social Sciences », *Annual Review of Sociology*, 2002, vol. 28, p. 167-195.

conceptuelles que les acteurs mobilisent pour catégoriser personnes, objets et pratiques) et frontières sociales (les écarts institutionnalisés de ressources et d'opportunités qui en résultent) aide à cadrer ce travail de codage. Les premières constituent une condition nécessaire mais non suffisante des secondes. Concrètement, lorsqu'une marque revendique un *savoir-faire artisanal* ou une *élégance parisienne*, elle trace une frontière conceptuelle. Que cette frontière se cristallise en frontière sociale, c'est-à-dire en écart effectif de prix et de positionnement marchand, dépend de son articulation avec d'autres dispositifs (réputation, canal de distribution, rareté organisée). Le discours institutionnel doit donc être pris comme un indice. Il ne suffit pas à prouver la position sociale d'une marque. Les analyses du chapitre 3 restent donc à ce niveau : elles comparent les registres discursifs avec les prix et les types d'espace commercial. Elles ne disent pas encore comment les consommateurs reçoivent ces distinctions.

1.2. Les outils statistiques actuels du commerce urbain

Les travaux quantitatifs sur le commerce urbain s'appuient sur quelques outils de la statistique publique. SIRENE est le répertoire administratif, tenu par l'Insee, qui recense les entreprises et leurs établissements et leur attribue un identifiant unique, le numéro SIREN pour l'entreprise et SIRET pour chaque établissement¹⁸. À chaque établissement est associé un code d'activité (APE), tiré de la NAF, la Nomenclature d'activités française, qui range chaque unité dans une classe d'activité économique et sur laquelle je reviens plus loin. À partir de ces données, l'Insee construit des pôles commerçants, moins un concept urbanistique préexistant qu'un objet produit par un algorithme de densité. Ces outils ont été développés pour des finalités précises, comptabilité nationale, suivi de l'emploi, aménagement commercial du territoire, auxquelles ils répondent bien. Mais ils n'ont pas été pensés pour saisir le positionnement social des espaces commerciaux, et c'est ce décalage qui m'intéresse.

La question posée ici est donc restreinte : ces catégories permettent-elles de saisir les différences sociales entre espaces commerciaux ? La méthode de l'Insee repose sur un critère de densité. Un pôle est détecté à partir d'un seuil de densité, autour d'au moins vingt établissements commerciaux proches les uns des autres, dans un voisinage de 200 mètres, puis

¹⁸ Insee, « Sirene » et « Nomenclature d'activités française », documentation en ligne, <https://www.insee.fr/fr/information/6675111> et <https://www.insee.fr/fr/information/2406147>, consulté le 28 mai 2026.

qualifié de centre-ville d'après la taille moyenne de ses établissements, sa proximité du centre communal et son environnement résidentiel, comme dans l'étude de Cazaubiel et Guymarc¹⁹ sur la déprise du commerce de proximité ; les pôles de périphérie relèvent de critères un peu différents²⁰. Ces critères sont morphologiques : ils décrivent la forme et la localisation du pôle, mais ne disent rien de son contenu. Un pôle dominé par des boutiques de luxe et un pôle dominé par des commerces de discount peuvent satisfaire les mêmes critères et être classés de la même manière.

Antoine Fleury, Hélène Mathian et Thérèse Saint-Julien²¹ adoptent une démarche différente. Ils mobilisent la BDCOM, base de données sur le commerce produite par l'Atelier parisien d'urbanisme à partir d'un relevé de terrain des locaux commerciaux parisiens. Son découpage est plus fin que celui de la NAF. À partir de cette source, les auteurs définissent la centralité commerciale par la rareté et la portée des commerces : un noyau est d'autant plus central qu'il concentre des commerces rares à large rayon d'attraction. Les deux approches ne donnent pas à voir la même ville. Un quartier riche en boulangeries et en coiffeurs est un pôle au sens de l'Insee sans être un noyau de centralité au sens de Fleury ; une rue de quelques galeries d'art peut l'être sans atteindre le seuil de densité de l'Insee.

La centralité commerciale est donc une construction méthodologique plutôt qu'une donnée, et chaque construction répond à une question légitime : couverture du territoire, dynamique des centres-villes, hiérarchie fonctionnelle de l'offre. Aucune ne porte cependant sur le positionnement social de l'offre. La rareté, chez Fleury et ses collègues, est un critère spatial, la diffusion inégale d'un type de commerce dans la ville, et non un critère social fondé sur le prix ou le public visé. Appliqués à la différenciation sociale du commerce, ces instruments en

¹⁹ Cazaubiel Arthur et Guymarc Gaël, « La déprise du commerce de proximité dans les centres-villes des villes de taille intermédiaire », *Insee Première*, 2019, n° 1782, <https://www.insee.fr/fr/statistiques/4248184>, consulté le 3 juin 2026.

²⁰ Bédoué Matthieu et Cohen Clément, « Le commerce de proximité : des pôles plus florissants en périphérie qu'en centre-ville », *Insee Première*, 2021, n° 1858, <https://www.insee.fr/fr/statistiques/5386414>, consulté le 3 juin 2026.

²¹ Fleury Antoine, Mathian Hélène et Saint-Julien Thérèse, « Définir les centralités commerciales au cœur d'une grande métropole : le cas de Paris intra-muros », *Cybergeo*, 2012.

renvoient une image fonctionnelle ou spatiale ; la hiérarchie sociale des enseignes leur échappe en grande partie par construction.

Ces études partent en outre du magasin individuel, ce qui rend difficile à saisir la logique d'ensemble d'un espace commercial, alors même qu'à cette échelle les effets d'agglomération deviennent décisifs et qu'un acteur (propriétaire, bailleur, gestionnaire) pèse sur la composition. L'opposition centre-ville / périphérie est elle-même morphologique : un pôle de périphérie dans les Yvelines et un en Seine-Saint-Denis apparaissent de la même façon.

Ces limites sont particulièrement visibles dans le cas des centres commerciaux, où la composition de l'offre est au moins en partie pilotée. Le réseau Westfield en Île-de-France l'illustre : Westfield Vélizy 2, dans les Yvelines, accueille des enseignes comme Sandro, Maje, COS ou Tommy Hilfiger, d'un segment intermédiaire à premium ; Westfield Rosny 2, en Seine-Saint-Denis, accueille davantage Kiabi, Gemo, La Halle ou Jennyfer, d'un segment plus accessible. À l'échelle de la NAF, une grande partie de ces enseignes peut pourtant relever du même code d'activité. La différence tient à la composition de l'offre plus qu'au nombre d'établissements : hiérarchie locale des marques, écarts de prix, position relative de chaque enseigne. Ces outils mesurent l'activité économique des établissements bien mieux que le positionnement social des enseignes.

1.3. La NAF : convention statistique et ses limites pour l'analyse sociale

La NAF, Nomenclature d'activités française, constitue l'un des socles du classement statistique des activités commerciales. Chaque entreprise, et chacun de ses établissements, se voit attribuer un code d'activité principale (APE) : 47.71Z pour le commerce de détail d'habillement en magasin spécialisé, 56.10A pour la restauration traditionnelle. Alignée sur les nomenclatures européenne et internationale, elle constitue une convention stable mais contrainte.

Pour comprendre ce qu'elle voit et ce qu'elle ne voit pas, il faut revenir à Desrosières²² : quantifier suppose d'abord de convenir, puis de mesurer. Une convention statistique est un accord sur les découpages qui rendent des objets comparables, et donc sur les différences qui resteront hors du cadre. La NAF classe les unités selon ce qu'elles produisent ou vendent ; le

²² Desrosières Alain, *La Politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte, 1993, notamment p. 9-16 et p. 397-418.

prix, le profil de la clientèle ou l'image projetée restent en dehors. Elle est cohérente avec ses usages d'origine, et sa limite n'apparaît que lorsqu'on la réutilise pour analyser la position sociale des enseignes.

Le code 47.71Z en donne l'exemple le plus net : il peut réunir des établissements aussi éloignés qu'une boutique de luxe et une enseigne populaire. Les cas les plus instructifs sont toutefois les positionnements intermédiaires. Des marques comme Sandro, Maje ou COS, plus chères que Zara mais en deçà du luxe, occupent un entre-deux difficile à coder : la frontière entre milieu de gamme et premium dépend du référentiel adopté et de la catégorie de produits. Un manteau Sandro à 350 euros et un sac COS à 120 euros ne se comparent pas facilement. C'est un problème de codage que la NAF n'a pas à affronter, puisqu'elle classe une activité. Dans mon enquête, le problème est plutôt de situer les enseignes les unes par rapport aux autres.

La difficulté s'accroît avec le lieu d'implantation. Une même enseigne, à gamme et prix comparables, n'occupe pas la même position relative selon son voisinage : entouré d'Hermès et de Saint Laurent, un magasin peut apparaître comme l'option la plus accessible ; entouré de Kiabi et de Gemo, comme l'une des plus hautes. Cette variation positionnelle, plus que la simple présence de la marque, justifie l'unité d'observation enseigne $\times$ centre retenue au chapitre 2.

Ces limites tiennent à la finalité de la NAF plus qu'à un défaut, et elles sont difficiles à lever. En 2018, à un député demandant la création d'un code pour les magasins bio, le gouvernement répondait qu'un code APE résume l'activité principale d'une entreprise et qu'introduire une sous-catégorie « bio » dans de nombreuses classes reviendrait à multiplier les catégories sans cohérence suffisante avec la logique d'activité de la nomenclature.²³ Le même raisonnement vaut pour le positionnement social des enseignes : empiriquement pertinent pour une enquête sociologique, il n'entre pas facilement dans la logique administrative de la NAF. C'est pourquoi il faut construire d'autres variables.

1.4. Construire de nouvelles variables : enjeux, hypothèses et limites

Si la NAF ne permet pas de saisir le positionnement social des enseignes, que faut-il faire ? La voie suivie dans ce mémoire est celle de la construction de nouvelles variables. Reprenant

²³ Question écrite n° 6258 de Frédérique Lardet et réponse du ministère de l'Économie et des Finances, Assemblée nationale, 15^e législature, 2018, en ligne : questions.assemblee-nationale.fr/q15/15-6258QE.htm.

Laurent Thévenot²⁴, on peut traiter la construction de catégories analytiques comme un investissement de forme. Ce qui est intéressant dans cette notion, c'est le mot investissement : produire une catégorie qui fonctionne, c'est-à-dire qui soit assez stable et assez partagée pour être utilisable, coûte cher. La NAF est un investissement massif, soutenu par des décennies de coordination. Les variables que je propose ici sont beaucoup plus modestes et plus discutables. Elles ont surtout un intérêt pratique : elles permettent de décrire des différences que la NAF ne montre pas.

Quelles variables faudrait-il pour saisir la différenciation sociale ? La plus évidente serait le segment tarifaire, mais la difficulté est immédiate. Les cas extrêmes (luxé d'un côté, discount de l'autre) se classent aisément ; les cas intermédiaires beaucoup moins, comme l'illustrent Sandro et COS. Le positionnement tarifaire ne se compare pas non plus entre catégories de produits : un restaurant à 40 euros et un vêtement à 200 euros n'occupent pas la même position dans leurs secteurs respectifs. Il faudrait construire le segment catégorie par catégorie, ce qui multiplie le travail.

À ce stade, le prix ne suffit plus. Le registre discursif des marques est une autre piste, et c'est celle que je privilégie dans ce mémoire. Cette piste s'appuie sur une intuition formulée par Pierre Bourdieu et Yvette Delsaut²⁵ dans leur article sur le couturier et sa griffe. Ils y montrent que la valeur d'un bien tient aussi à un travail symbolique (signature, récit, mise en scène) qui n'est pas réductible à ses propriétés matérielles. Une griffe n'est pas une simple étiquette de marque : c'est un dispositif de production de la croyance dans la valeur de l'objet. Cette intuition donne un appui théorique au fait de traiter le discours de marque comme une trace du travail symbolique de valorisation. En analysant ce que les marques disent d'elles-mêmes sur leurs sites institutionnels, on accède à une forme de mise en récit de cette valeur plutôt qu'à la valeur elle-même. Le cadre de Boltanski et Esquerre ne se transpose pas directement à mon corpus : je ne retrouve pas simplement leurs quatre formes dans les sites de marques. Leur travail m'aide surtout à prendre au sérieux les récits de patrimoine, de rareté ou de création comme des éléments de valorisation, même si les catégories que je retiens (luxé patrimonial, création d'auteur, standard) restent propres à l'enquête. Mais coder un registre discursif n'est

²⁴ Thévenot Laurent, « Les investissements de forme », dans Thévenot Laurent (dir.), *Conventions économiques*, Paris, Presses universitaires de France, 1986, p. 21-71.

²⁵ Bourdieu Pierre et Delsaut Yvette, « Le couturier et sa griffe : contribution à une théorie de la magie », *Actes de la recherche en sciences sociales*, 1975, vol. 1, n° 1, p. 7-36.

pas trivial. Les sites de marques internationales sont souvent en anglais, la communication varie selon les pays, et les catégories de discours (lifestyle, tradition, savoir-faire) ne sont pas forcément stables dans le temps. Une enseigne peut changer de registre en quelques années sans changer de segment tarifaire. Il faut donc d'abord convenir de ce qu'on veut voir avant de pouvoir le mesurer, et cette convention initiale est loin d'être évidente.

L'appartenance à un groupe d'entreprise (LVMH, Kering, Inditex, Estée Lauder ou absence de rattachement identifié) a été renseignée partiellement, comme variable de contrôle ; cette information n'étant pas disponible dans SIRENE et devant être reconstituée manuellement, sa couverture limitée empêche d'en faire un axe central de l'analyse.

À ces variables sociales s'ajoutent les variables spatiales que je construis sur la base d'une typologie sociologiquement informée des espaces commerciaux étudiés. Quatre types sont distingués : grand magasin central parisien, centre commercial de couronne, centre commercial généraliste, espace de commerce de transit. Cette typologie ne reprend pas telle quelle les nomenclatures professionnelles du Conseil national des centres commerciaux ou de Procos, qui s'appuient principalement sur des critères de surface, de localisation ou de rayonnement commercial. Elle combine plutôt trois dimensions : un format commercial, un positionnement marchand et une fonction d'achat.

Les variables nouvelles mobilisées dans le mémoire sont récapitulées en annexe, avec leur source, leur fonction analytique et leur limite principale. Elles combinent des variables discursives (forme dominante, intensité d'enrichissement, registres lexicaux), des variables de prix (médian, maximal, amplitude, position sectorielle), des variables spatiales (type d'espace, position enseigne $\times$ centre) et des variables de structure (secteur, cohorte de fondation, appartenance à un groupe). Ces variables ne sont pas indépendantes : la forme discursive dominante, l'intensité d'enrichissement et les registres lexicaux se recoupent partiellement, parce que les marques mobilisent souvent plusieurs ressources symboliques à la fois.

Ces variables déplacent l'observation : elles font apparaître les prix, les registres discursifs, les voisinages et les positions relatives des enseignes que les catégories ordinaires laissent hors champ.

Trois hypothèses guident la lecture des résultats. La première porte sur la dimension typologique. Je suppose que les formes discursives dominantes construites à partir du cadre de l'enrichissement se distribuent dans le corpus de manière différenciée et que cette distribution

est associée à des écarts systématiques de prix. La seconde porte sur la grammaire interne du discours. Je suppose que les différents registres mobilisés entretiennent avec le prix des associations contrastées : les registres patrimoniaux et de luxe explicite seraient positivement associés au prix, tandis que les registres du care et de la responsabilité relèveraient peut-être d'une autre logique de valorisation. La troisième est une hypothèse de sélection spatiale : les types d'espace commercial sélectionneraient des marques différentes plus qu'ils ne renchériraient les mêmes discours.

Chapitre 2. Construction et exploration des données

2.1. Le terrain : neuf espaces commerciaux franciliens

La base n'a pas existé d'un seul coup. Elle s'est construite par ajouts, corrections, recodages et renoncements. Le terrain d'étude se compose de l'ensemble des enseignes recensées dans neuf espaces commerciaux situés en Île-de-France. Le choix de ces neuf espaces relève d'une construction de recherche raisonnée plutôt que d'un échantillonnage représentatif. Trois grands magasins centraux parisiens ont été retenus : Le Bon Marché Rive Gauche, La Samaritaine et le BHV Rivoli. Quatre espaces commerciaux de couronne complètent la liste : Westfield Les 4 Temps à La Défense, Westfield Vélizy 2, BHV Parly 2 et Westfield Rosny 2. Un centre commercial généraliste a également été inclus : Italie 2 dans le treizième arrondissement de Paris.²⁶ Enfin, un espace de commerce de transit, Westfield Aéroville, situé près de Roissy-Charles-de-Gaulle, complète l'échantillon.

Le terrain est volontairement situé : il porte sur neuf espaces franciliens et n'autorise aucune inférence sur le commerce français dans son ensemble. Le biais principal est assez évident : Paris occupe une position centrale dans le paysage commercial national et concentre une part disproportionnée des maisons patrimoniales, des grands magasins historiques et des marques de luxe françaises. Cette concentration peut servir l'analyse, dans la mesure où elle met à disposition une densité de cas suffisante pour observer finement les opérations discursives. Le périmètre effectif ne couvre toutefois qu'une partie très située de la région parisienne. Les

²⁶ BHV Parly 2 est traité ici comme un espace commercial de couronne, même s'il relève plutôt du format grand magasin inséré dans un centre commercial.

neuf espaces retenus, quatre à Paris intra-muros et cinq en couronne, ne représentent ni l'ensemble de l'offre francilienne ni l'ensemble de l'offre française.

La nomenclature retenue (grand magasin central, espace commercial de couronne, centre commercial généraliste, espace de commerce de transit) ne reprend telle quelle aucune typologie professionnelle existante. Les nomenclatures issues du Conseil national des centres commerciaux, et les classements produits par des acteurs comme Procos, s'appuient surtout sur des critères de taille, de rayonnement, de localisation ou de performance commerciale. Ces catégories sont utiles pour les bailleurs, les investisseurs ou les professionnels de l'immobilier commercial. Elles saisissent une réalité économique importante, mais ce n'est pas exactement celle qui m'intéresse ici. Le contraste entre grand magasin patrimonial parisien et centre commercial généraliste ne se réduit pas à une opposition de taille. La typologie adoptée recompose donc trois dimensions : un format commercial (grand magasin historique, centre commercial, espace de transit), un positionnement marchand (premium, généraliste) et une fonction d'achat (délibérée, contingente). C'est une typologie construite pour l'enquête. Elle n'a pas la prétention d'être une typologie professionnelle générale.

La base finale rassemble 2 978 observations enseigne × centre, correspondant à 2 173 marques uniques incluses dans la base. La distribution des observations entre les quatre types d'espace est fortement asymétrique. Les grands magasins centraux concentrent environ 68 % des observations, les espaces commerciaux de couronne 24 %, l'espace de commerce de transit 4 %, et les centres commerciaux généralistes 3,3 %. Cette asymétrie limite la puissance statistique des analyses portant sur les espaces les moins représentés, surtout les centres généralistes. Une partie du déséquilibre tient au périmètre choisi : les espaces patrimoniaux et les grands magasins centraux rassemblent davantage d'enseignes dont le discours institutionnel est long et exploitable. Mais une autre partie tient aussi à la nature des matériaux disponibles : dans les centres les moins dotés en marques patrimoniales, et surtout à Westfield Rosny 2 et à Italie 2, une part plus importante des enseignes relève du mass-market, de services ou de commerces plus locaux, pour lesquels le discours institutionnel est parfois plus court, plus fonctionnel ou plus difficile à stabiliser comme matériau comparable. Les inférences sur ces centres peu fournis demeurent donc fragiles et sont encadrées de réserves explicites tout au long de l'exposé des résultats.

2.2. Sources et collecte

La construction de la base a mobilisé trois familles de sources. La base a été construite en plusieurs étapes : d’abord les annuaires des centres, puis les sites de marques, les catalogues publics, SIRENE et enfin plusieurs recodages. Chaque étape rend les données un peu plus comparables, mais elle simplifie aussi le matériau de départ.

La première famille rassemble les répertoires de marques publiés par les centres commerciaux sur leurs sites institutionnels. Chaque centre maintient un annuaire des enseignes présentes, généralement assorti d’une catégorisation par segments. Ces répertoires ont été extraits par scripts d’extraction automatisée écrits en Python, à raison d’un script par centre.

La deuxième famille rassemble les sites institutionnels des marques elles-mêmes. Pour chaque marque identifiée, l’URL du site officiel a été recherchée. Une procédure d’extraction structurée a permis de récupérer le discours présenté sur les pages : À propos, *Notre histoire*, *La maison*, *Heritage*, *Manifeste*, *Nos engagements*, ou leurs équivalents en anglais. Le texte ainsi collecté a été normalisé, débarrassé des éléments de navigation et de publicité, puis stocké dans une colonne unifiée nommée `discourse_canonical`. Cette extraction n’a pas toujours été nette : pour certaines enseignes, le site officiel séparait mal le discours institutionnel du discours commercial courant, et il a fallu trancher au cas par cas. Lorsque c’était possible, des prix ont également été relevés à partir des catalogues publics des marques. Pour chaque marque, jusqu’à dix prix de produits ont été conservés. Ces dix prix correspondent à ceux que j’ai pu relever dans les catalogues publics au moment de la collecte, en évitant les doublons manifestes et les produits qui ne relevaient pas directement de l’offre principale de la marque. Ce n’est donc pas un tirage aléatoire. Ce choix ne produit pas un panier homogène de biens strictement comparables ; il permet plutôt de construire une petite distribution tarifaire, à partir de laquelle sont calculés trois indicateurs : un prix d’entrée, un prix médian et un prix maximal.

Ce choix vient d’un problème concret rencontré pendant la collecte. Une même marque peut proposer des produits très hétérogènes : Nike, par exemple, vend aussi bien des chaussettes à une dizaine d’euros que des chaussures à plus de cent euros, voire des pièces de collaboration nettement plus chères. Retenir un seul prix par marque aurait donc produit un indicateur trompeur. Le prix d’entrée renseigne le seuil minimal d’accès à l’offre observée, le prix médian donne une indication du niveau courant des produits relevés, tandis que le prix

maximal permet d'approcher le plafond visible de la gamme. Cette méthode ne garantit pas une comparabilité stricte entre produits ou entre marques. Elle vise plutôt à rendre visible l'étendue de l'offre tarifaire disponible publiquement. Lorsque les prix n'étaient pas publics, trop instables ou difficilement récupérables de manière fiable, ils ont été laissés manquants.

La troisième famille de sources est administrative. Les informations issues de SIRENE ont été recherchées pour chaque marque, à partir de son nom commercial, parfois complété par l'adresse ou le centre d'implantation, afin de récupérer le numéro SIREN, la dénomination légale, le code NAF et son libellé, ainsi que la date de création de l'entité juridique. Le rattachement n'allait pas toujours de soi : lorsque le nom commercial différait de la dénomination légale, ou qu'une même enseigne renvoyait à plusieurs entités, l'appariement a dû être vérifié à la main. Cette interrogation a été automatisée via l'API publique recherche-entreprises.api.gouv.fr. SIRENE ne fournit pas directement une information exploitable sur l'appartenance des enseignes à un groupe d'entreprises. Cette variable a donc été ajoutée séparément, à partir de recherches complémentaires sur les marques et leurs sociétés mères, et reste renseignée de manière partielle.

L'ensemble de la collecte s'est déroulé entre janvier et mai 2026. Les prix observés correspondent à un état du marché à cette période. La mise à jour des catalogues, l'évolution des collections saisonnières et les éventuelles révisions tarifaires postérieures à cette date ne sont pas prises en compte. C'est une limite que partage toute enquête transversale.

2.3. Le codage du discours : trois schèmes

Le codage a été séparé en trois niveaux : la forme dominante du discours, l'intensité d'enrichissement et les registres lexicaux activés. Coder la seule forme dominante écrasait les cas intermédiaires, où une marque reste standard dans son architecture générale tout en mobilisant plusieurs marqueurs d'ancienneté ou de savoir-faire ; coder les seuls registres lexicaux aurait effacé la différence entre une mention ponctuelle et un récit entièrement organisé autour du patrimoine. Séparer les trois niveaux évite de tout rabattre sur une seule variable.

Le premier schème, la forme dominante, est volontairement exigeant : une marque n'est codée `luxe_patrimoine` ou `creation_auteur` que si l'enrichissement organise l'ensemble de son discours. Or beaucoup de marques mobilisent ponctuellement un registre d'enrichissement comme un savoir-faire revendiqué ou une date de fondation. Leur discours

n'en reste pas moins commercial dans son architecture, et tourné vers l'accessibilité ou le prix bas. Les ranger simplement parmi les `standard` ferait perdre cette information. C'est pourquoi un deuxième schème, l'intensité, mesure le degré d'enrichissement indépendamment de la forme : il distingue, parmi les marques `standard`, celles qui n'enrichissent pas du tout leur discours de celles qui l'enrichissent partiellement ou fortement. Mais l'intensité reste une mesure agrégée : elle indique combien, pas quoi. Le troisième schème code donc séparément chaque registre mobilisé, ce qui permet de voir quelles grammaires un même discours combine et, surtout, de construire l'espace des registres analysé par l'analyse des correspondances multiples. Dans la pratique du codage, cette distinction a surtout servi à éviter deux erreurs : classer comme patrimoniale toute marque qui mentionne une date ancienne, ou perdre dans une grande catégorie `standard` les emprunts ponctuels à l'enrichissement. Aucun schème ne se déduit entièrement des deux autres, et c'est leur croisement que les analyses cherchent à exploiter.

Le premier schème distingue trois formes dominantes. La forme `standard` regroupe les discours dont la logique reste commerciale et descriptive : présentation du produit, du point de vente, des fonctionnalités, du positionnement de prix. La forme `luxe_patrimoine` regroupe les discours organisés autour de la valorisation patrimoniale : héritage, date ancienne, savoir-faire artisanal, transmission générationnelle, revendication explicite du luxe. La forme `creation_auteur` regroupe les discours organisés autour de la revendication d'une signature : designer nommé, studio créatif, manifeste, univers singulier, vision esthétique articulée. Cette dernière forme s'éclaire par référence à l'article de Bourdieu et Delsaut²⁷ sur le couturier et sa griffe, où les auteurs montrent que la valeur d'un bien tient aussi à un travail de production symbolique : la signature, la mise en récit, le rattachement à une figure d'auteur. La forme `creation_auteur` code précisément ce dispositif lorsqu'il devient l'organisateur principal du discours de la marque.

Le codage du premier schème a mobilisé ActiveTigger, un outil d'annotation assistée par apprentissage actif développé autour de Julien Boelaert, Étienne Ollion et leurs

²⁷ Bourdieu Pierre et Delsaut Yvette, « Le couturier et sa griffe : contribution à une théorie de la magie », *Actes de la recherche en sciences sociales*, 1975, vol. 1, n° 1, p. 7-36.

collaborateurs²⁸. Il propose itérativement à l’annotateur humain les cas les plus informatifs à coder. Un total de 351 documents, choisis comme cas informatifs ou paradigmatiques, a été annoté manuellement. Sur cette base, un classifieur sklearn de régression logistique a été entraîné, puis appliqué à l’ensemble du corpus discursif. La performance du classifieur, mesurée par macro F1 sur validation croisée à cinq plis, atteint 0,774. Ce score constitue un indicateur de qualité globale du classement, mais il ne garantit pas l’absence d’erreurs de classification. Les erreurs probables concernent surtout les cas intermédiaires, par exemple les marques standard qui mobilisent fortement le patrimoine sans être des maisons de luxe. On peut rapprocher cette démarche, plus modestement, de ce que Laura Nelson appelle *computational grounded theory*²⁹, qui décrit assez bien le pipeline : annotation interprétative sur cas paradigmatiques, extension par classifieur, puis, dans mon cas, validation partielle par croisement avec des variables externes.

Le deuxième schème, qui code l’intensité d’enrichissement, repose sur une règle déterministe. Si la forme dominante est `luxe_patrimoine` ou `creation_auteur`, l’intensité est codée `fort` par construction, puisque l’enrichissement est constitutif de la forme dominante. Si la forme dominante est `standard`, l’intensité dépend du nombre de marqueurs détectés et du nombre de registres mobilisés. Un discours standard sans marqueur détecté est codé `non`. Un discours standard mobilisant un à trois marqueurs ou un à deux registres est codé `faible`. Un discours standard mobilisant quatre marqueurs et plus, ou trois registres et plus, est codé `fort`.

Le troisième schème code la mobilisation de registres lexicaux. Il comprend d’abord sept registres principaux, dont les marqueurs reposent sur des dictionnaires structurés. Chaque registre est défini par une liste de chaînes de caractères et de motifs d’expressions régulières. Ces dictionnaires sont les mêmes que ceux qui servent à calculer l’intensité au deuxième schème ; seul leur usage diffère. L’intensité agrège les marqueurs détectés en un degré global (combien de registres, combien de marqueurs) ; le troisième schème, lui, conserve chaque

²⁸ ActiveTigger, outil open-source d’annotation et de classification de textes en sciences sociales computationnelles, développé autour de Julien Boelaert, Étienne Ollion et leurs collaborateurs ; voir la documentation du projet sur activetigger.com.

²⁹ Nelson Laura K., « Computational Grounded Theory: A Methodological Framework », *Sociological Methods & Research*, 2020, vol. 49, n° 1, p. 3-42.

registre comme une variable binaire distincte (présence ou absence), directement mobilisable comme variable active dans l'analyse factorielle. La même détection lexicale sert donc, dans un cas, à mesurer une densité, dans l'autre, à cartographier des contenus, sans que les deux schèmes fassent double emploi. Les sept registres sont : patrimoine, auteur-design, tendance-lifestyle, care-émotion, responsabilité-éthique, rareté-collection, expression-identité. Les dictionnaires ont été construits dans une démarche hybride combinant codage déductif et codage inductif. Les registres ont été définis a priori à partir du cadre théorique de Boltanski et Esquerre, et leurs marqueurs lexicaux ont été enrichis itérativement par lecture des cas empiriques. Cinq registres complémentaires plus stricts s'ajoutent à ce dispositif et précisent la grammaire de l'enrichissement classique : ancienneté mobilisée, savoir-faire, origine franco-parisienne, archive intemporelle, luxe explicite. Au total, douze registres sont ainsi codés, sept principaux et cinq complémentaires. Ces registres ne couvrent pas toutes les ressources symboliques possibles ; ils correspondent aux grammaires qui revenaient le plus souvent dans le corpus après plusieurs lectures, lues à partir du cadre de Boltanski et Esquerre. Leurs frontières sont définies opérationnellement par les dictionnaires de marqueurs, à raison de quinze à trente motifs par registre, détaillés dans le codebook joint ; un même discours peut activer plusieurs registres à la fois, ce qui distingue ce schème, à étiquettes multiples, des deux premiers, à étiquette unique. Le registre rareté-collection, présent dans moins de 1 % des discours, est trop rare pour être stable dans les analyses multivariées : il est conservé dans la base descriptive mais écarté des analyses factorielles et des régressions, qui portent donc sur onze registres actifs.

Le pipeline analytique implémente l'ensemble du codage en R, à l'exception du classifieur de forme qui reste produit par Python. Pour les sept registres lexicaux et les cinq registres complémentaires, l'extraction par regex est réalisée en R.

2.4. Variables analytiques

Plusieurs variables structurelles ont été construites pour servir de support à l'analyse.

Le sous-groupe combine la forme dominante et l'intensité d'enrichissement en cinq modalités principales : `standard_non`, `standard_faible`, `standard_hybride`, `creation_auteur`, `luxe_patrimoine`. La modalité `standard_hybride` mérite une précision méthodologique explicite. Cette catégorie n'est pas conçue comme une variable indépendante des registres discursifs. Elle est construite pour identifier, au sein des marques

standard, les cas de combinaison forte de plusieurs registres. Guy Degrenne a constitué l'un des cas-limites du codage. Le discours active fortement l'ancienneté, le patrimoine et le savoir-faire, mais il ne fonctionne pas comme celui d'une maison de luxe : il reste organisé autour d'une marque grand public d'arts de la table. Le classer en `luxe_patrimoine` aurait surcodé l'enrichissement ; le laisser dans standard aurait fait disparaître une partie du travail symbolique. C'est ce type de cas qui justifie la modalité `standard_hybride`. Elle est mobilisée pour examiner si cette hybridation construite se trouve associée à des espaces commerciaux et à des niveaux de prix particuliers. Le test repose ainsi sur des variables externes (type d'espace, prix médian, prix maximal), qui ne participent pas à la construction du sous-groupe. Dans l'ACM, le sous-groupe comme les variables externes seront projetés comme variables supplémentaires.

Le secteur analytique a été construit par une cascade hiérarchique, combinant cinq sources ou opérations successives, pour pallier les lacunes du codage natif. La catégorie commerciale native fournie par les annuaires des centres n'est disponible que pour quatre des neuf centres. Une procédure de propagation par `brand_key` permet de récupérer la catégorie d'une marque vue ailleurs. Lorsque la propagation est insuffisante, la cascade mobilise successivement le libellé NAF de l'Insee, la catégorie produit native de la marque, et en dernier recours le code NAF lui-même. Cette cascade réduit le taux de non-classification de 55 % à environ 0,3 %. Les onze secteurs finaux sont : beauté, mode, maison, maroquinerie-chaussure, joaillerie-horlogerie, enfant, lingerie, bricolage, services spécialisés, autre et non classé.

L'âge des marques a été reconstruit par une cascade similaire en plusieurs sources, qui combine successivement l'année de fondation objective collectée, l'année revendiquée par la marque, l'année extraite par regex de son discours, et en dernier recours la date de création de l'entité juridique enregistrée à l'Insee. Cette cascade réduit le taux d'inconnu de 80 % à environ 2 %. Ce recours à l'Insee pose toutefois un problème. La date d'immatriculation désigne l'entité juridique, pas nécessairement la marque elle-même. Pour environ 12 % des observations, cette substitution tend à rajeunir artificiellement certaines marques, en les rattachant à la date récente de leur entité juridique plutôt qu'à leur fondation historique. J'en tiens donc compte lorsque j'interprète les cohortes de fondation, en particulier pour les marques anciennes rattachées à une entité juridique récente.

Les indicateurs de prix sont au nombre de cinq : prix médian par marque, prix maximal par marque, prix au 90^e percentile, prix relatif au secteur, amplitude max/min. Le choix de

l'indicateur principal a été difficile. La distribution des prix par marque suit une structure asymétrique. Le panier de chaque marque rassemble jusqu'à dix prix observables dans le catalogue public. Il forme une petite distribution tarifaire plutôt qu'un panier homogène de produits strictement comparables, et permet d'observer un point d'entrée, un niveau médian et un plafond de gamme. Pour les marques de luxe à gamme étendue, le panier mêle des articles d'entrée de gamme (parfums, cosmétiques, accessoires) et des pièces signature de catégorie haute (sacs, joaillerie). La médiane se trouve alors tirée vers le bas, alors même que le plafond reste très élevé. Le prix médian rend les marques comparables, mais il écrase le luxe étendu lorsque le panier contient des parfums ou de petits accessoires : un sac ou une pièce de joaillerie fixe un plafond symbolique que la médiane efface. J'ai donc retenu le prix maximal observé comme indicateur de plafond de gamme, en contrôlant chaque résultat par le 90^e percentile, dont le modèle, présenté en annexe de robustesse, conduit aux mêmes conclusions. Le prix maximal observé est donc traité comme un plafond de gamme dans le panier collecté, et non comme une mesure du niveau moyen de la marque. Le prix relatif au secteur est calculé comme un écart logarithmique à la médiane sectorielle pour neutraliser les hétérogénéités sectorielles. Les modalités « non classé » et « autre » du secteur, qui ne disposent pas d'un benchmark sectoriel fiable, sont exclues du calcul de cette dernière variable.

2.5. Limites de la production des données

La construction des données mobilisée dans ce mémoire combine plusieurs sources hétérogènes dont l'articulation pose des questions de validité qu'il faut expliciter.

Les limites de la base ne sont pas toutes du même ordre. J'en distingue trois. La première tient au codage interprétatif. La forme discursive dominante a été codée manuellement sur 351 cas. L'annotation a été réalisée par une seule personne, l'auteur du mémoire. Aucun calcul d'accord inter-codeurs n'a pu être conduit, ce qui constitue une limite reconnue. La validité du codage repose donc surtout sur la stabilité de mes décisions et sur le codebook utilisé pour les expliciter. ActiveTigger a tout de même structuré l'annotation : l'outil m'a régulièrement renvoyé vers les cas sur lesquels le classifieur hésitait. Cela m'a obligé à travailler les frontières entre catégories, mais cela ne remplace pas un accord inter-codeurs. Dans un sens proche de ce que Lahire³⁰ montre à propos des assemblages culturels individuels, il faut aussi

³⁰ Lahire Bernard, *La Culture des individus. Dissonances culturelles et distinction de soi*, Paris, La Découverte, 2006, p. 141-146.

reconnaître que les marques présentent souvent des assemblages composites qui résistent à toute typologie pure. Les cas codés comme `standard_hybride` en sont l'illustration : ils ne cadrent ni avec la forme `standard` ni avec les formes `patrimoniaire` ou `auteur`, et c'est précisément leur position intermédiaire qui les rend analytiquement intéressants.

La deuxième limite vient des dictionnaires. Ils ratent certaines formulations atypiques et, à l'inverse, ils comptent parfois un mot comme marqueur alors que le contexte ne correspond pas vraiment au registre visé. C'est la limite d'un codage lexical : il stabilise la comparaison au prix d'une partie de la nuance des textes.

La dernière limite est plus matérielle : les prix. Les dix prix relevés ne décrivent pas toujours les mêmes types de produits selon les marques. Les comparaisons portent donc sur des gammes visibles, pas sur des paniers strictement équivalents. C'est pour cette raison que je lis ensemble le prix médian, le prix maximal, le p90 et l'amplitude, plutôt qu'un indicateur isolé. Pour les marques dont les prix ne sont pas publiquement accessibles, en particulier dans le segment du luxe haut de gamme, les prix sont laissés manquants ou traités comme non comparables lorsque aucune source publique fiable ne permet de les établir.

Les sources mobilisées sont exclusivement publiques (annuaires des centres, sites institutionnels des marques, base SIRENE) et ne portent pas sur des données personnelles ; l'extraction a été conduite à un rythme limité, pour un usage strictement académique. L'ensemble des analyses peut être reproduit à partir d'un script R documenté (chargement, extraction lexicale, construction des variables, tableaux, ACM et régressions), la seule étape située en amont étant la prédiction de la forme discursive dominante par un classifieur Python.

Chapitre 3. L'enrichissement à l'épreuve des prix : typologie, grammaire et sélection spatiale

3.1. Vue d'ensemble du corpus

Une fois la base stabilisée, il restait à voir si les variables construites faisaient apparaître autre chose qu'un simple effet de secteur ou de luxe. La base finale rassemble 2 978 observations $\text{enseigne} \times \text{centre}$, correspondant à 2 173 marques uniques incluses dans la base. Parmi ces observations, 2 531 disposent d'un discours suffisamment exploitable pour entrer dans l'ACM ; les autres restent dans la base mais ne participent pas à l'analyse factorielle des

registres. La distribution des sous-groupes issus du codage discursif est résumée dans le tableau suivant.

Sous-groupe	N (obs.)	%
standard_non	919	30,9
standard_faible	716	24,0
standard_hybride	121	4,1
creation_auteur	472	15,8
lux_e_patrimoine	305	10,2
autre	445	14,9
Total	2 978	100

La modalité autre est une catégorie résiduelle, qui regroupe les cas non classés dans les cinq sous-groupes interprétables. Bien que résiduelle, elle se comporte de façon homogène du côté des prix, avec la médiane la plus basse du corpus (32 €) : les discours difficiles à classer relèvent surtout de marques de mass-market ou de services. La catégorie standard_hybride reste minoritaire. Les croisements qui la concernent sont sensibles aux faibles effectifs ; leur intérêt tient moins à leur généralisation qu'à la position intermédiaire qu'ils rendent visible.

La distribution des observations selon trois dimensions (tableau détaillé en annexe, N = 2 978 observations) est fortement asymétrique pour le type d'espace (grand magasin central 68,0 %, espace de couronne 24,3 %, transit 4,4 %, généraliste 3,3 %), plus équilibrée pour le secteur (mode 20,2 %, services spécialisés 15,2 % et beauté 14,1 % en tête), et nettement concentrée sur les marques contemporaines pour la cohorte de fondation (42,6 % des observations rattachées à des marques fondées après 2000, contre 9,9 % avant 1900).

La prévalence des onze registres actifs codés dans le corpus présente une asymétrie marquée, résumée dans le tableau suivant.

Registre	Prévalence (% des observations)
Patrimoine	40,5 %
Ancienneté	34,0 %

Care-émotion	22,0 %
Tendance-lifestyle	19,0 %
Auteur-design	11,8 %
Savoir-faire	11,0 %
Responsabilité-éthique	8,2 %
Luxe explicite	7,5 %
Expression-identité	7,4 %
Archive intemporelle	6,9 %
Origine franco-parisienne	2,7 %

Ces pourcentages sont calculés au niveau des observations enseigne × centre, et non au niveau des marques uniques : une marque présente dans plusieurs centres pèse donc plusieurs fois dans la distribution, ce qui est cohérent avec l'unité d'observation retenue dans le mémoire. Les registres patrimoniaux et temporels (patrimoine, ancienneté) sont nettement les plus prévalents, tandis que des registres comme l'origine franco-parisienne ou l'archive intemporelle restent marginaux. Cette distribution prépare la lecture de l'ACM. Les registres fréquents, comme le patrimoine ou l'ancienneté, structurent fortement le corpus ; des registres plus rares peuvent aussi peser lorsqu'ils sont concentrés sur quelques profils très spécifiques.

La distribution des prix par sous-groupe, résumée dans le tableau suivant, permet une première lecture de l'association entre forme discursive et niveau de prix.

Sous-groupe	Médiane des prix médians (€)	Médiane du prix maximal (€)	Q3 du prix maximal (€)	Médiane de l'amplitude max/min
standard_non	46	119	282	4,8
standard_faible	87	169	430	4,2
standard_hybride	124	220	496	4,0
creation_auteur	225	390	1 200	5,0
luxe_patrimoine	125	340	1 350	5,6
autre	32	69	230	4,4

La progression de la médiane des prix médians est nette, du sous-groupe autre (32 €) au sous-groupe `creation_auteur` (225 €), mais elle ne raconte pas toute l'histoire. La médiane, seule, ne suffisait pas. Elle sous-estime surtout le positionnement du sous-groupe `lux_e_patrimoine`, dont le panier peut mêler des produits d'entrée de gamme et des pièces beaucoup plus chères : sa médiane (125 €) est proche de celle de `standard_hybride` (124 €), alors que son plafond est bien plus élevé, comme le montrent la médiane du prix maximal (340 €) et surtout le troisième quartile du prix maximal (1 350 €). L'amplitude max/min du panier indique encore plus nettement la singularité du sous-groupe `lux_e_patrimoine` : 5,6, contre 4,0 pour `standard_hybride` et entre 4,2 et 4,8 pour les autres sous-groupes standard. Cette amplitude élevée peut se lire comme un indice d'une opération marchande caractéristique du luxe étendu : dans le panier observé, des articles d'entrée de gamme (parfums, accessoires) coexistent avec des pièces signature de catégorie haute (sacs, joaillerie). Certaines marques rendent ainsi visibles des objets relativement accessibles tout en maintenant un plafond de prestige élevé.

3.2. L'espace factoriel des registres discursifs

L'analyse des correspondances multiples a été réalisée sur l'échantillon des observations dont le discours est utilisable, soit 2 531 enseigne $\times$ centre. Les onze registres actifs ont été codés en variables binaires (0 = absence, 1 = présence). Plusieurs variables supplémentaires ont été projetées sans participer à la construction de l'espace : le sous-groupe, le type d'espace commercial, le secteur analytique, la cohorte de fondation, et les indicateurs de prix catégorisés. Les deux premières dimensions capturent 18,2 % et 11,4 % de l'inertie brute, soit 29,6 % cumulés (scree plot reporté en annexe). Ces taux bruts sous-estiment toutefois systématiquement la part de structure, propriété géométrique connue de l'ACM où le grand nombre de modalités gonfle l'inertie totale ; les taux modifiés de Benzécri portent la première dimension à 90,4 % et les deux premières à 96,1 % (quatre axes seulement ont une valeur propre supérieure à 1/Q). La première dimension concentre donc l'essentiel de la structure ; je conserve la deuxième comme opposition secondaire, en tenant compte de sa plus faible inertie modifiée. Le discours étant codé au niveau de la marque, une enseigne présente dans plusieurs centres y répète exactement le même profil de registres : les 2 531 observations ne correspondent donc pas à 2 531 textes indépendants, et les marques multi-centres pèsent davantage dans la décomposition de l'inertie. Une ACM recalculée sur les marques dédoublées (une ligne par marque, soit 1 875 marques) laisse cette décomposition presque

inchangée — 17,8 % et 11,6 % pour les deux premières dimensions — et reproduit la première dimension quasiment à l'identique (corrélation des coordonnées de modalités proche de 0,98) ; la deuxième dimension, d'inertie voisine de la troisième, se redéfinit dans un plan qui reste lui-même stable (Annexe B). La structure de l'espace discursif ne tient donc pas à la multiplicité des observations. Le couple enseigne $\times$ centre n'apporte d'ailleurs presque rien à cette analyse discursive, qui pourrait être menée au niveau des marques ; sa valeur tient à l'analyse des prix relatifs et de la position spatiale (section 3.6), où une même marque varie selon le centre.

Dans mes données, la première dimension oppose surtout les discours sans marqueurs patrimoniaux aux discours qui accumulent ancienneté, patrimoine, savoir-faire et luxe explicite (Figure 1). Je la nomme donc axe d'enrichissement patrimonial et artisanal. Du côté négatif se trouvent les modalités de non-mobilisation des registres patrimoine, ancienneté, archive, savoir-faire et luxe explicite. Du côté positif se trouvent les modalités de mobilisation de ces mêmes registres, auxquelles s'ajoute la revendication d'origine franco-parisienne. Cette première opposition factorielle oppose surtout l'accumulation de marqueurs temporels, artisanaux et institutionnels à leur absence. Elle ne sépare donc pas simplement le luxe des autres marques. C'est précisément là que le codage devient utile. Le rapprochement avec la forme collection de Boltanski et Esquerre — la valeur produite par l'inscription narrative dans un ensemble historiquement situé — éclaire cette opération, même si mes indicateurs restent beaucoup plus simples que leur modèle.

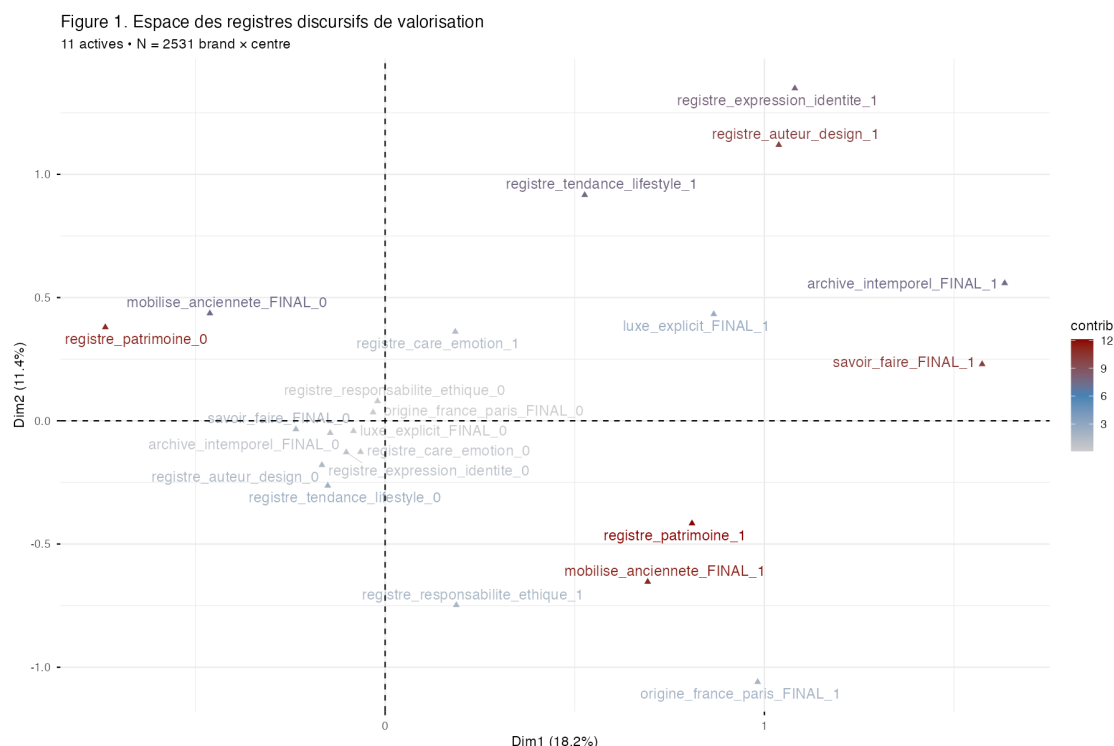


Figure 1. Espace factoriel des onze registres discursifs actifs. La coloration indique la contribution de chaque modalité aux deux premières dimensions.

La deuxième dimension n'est pas un simple prolongement de la première : elle sépare deux manières d'enrichir le discours, que je résume comme un axe d'expression créative face à un ancrage patrimonial. Du côté supérieur se trouvent les modalités de mobilisation des registres expression-identité, auteur-design, tendance-lifestyle et care-émotion ; le savoir-faire, l'archive et le luxe explicite se rapprochent aussi de ce pôle. Du côté inférieur se trouvent surtout les modalités patrimoine, ancienneté, origine franco-parisienne et responsabilité-éthique. Cette opposition distingue deux manières de construire une valeur symbolique enrichie : d'un côté une grammaire plus expressive, stylistique et créative ; de l'autre une grammaire d'ancrage, fondée sur le temps long, l'origine et l'inscription patrimoniale. Le deuxième axe reste plus fragile à interpréter, notamment à cause de la position du savoir-faire, située dans la partie supérieure du plan factoriel. Au total, l'existence d'une seconde dimension distincte de la première indique que la valorisation symbolique des marques ne se réduit pas à un continuum d'enrichissement croissant. Plusieurs grammaires de valorisation symbolique coexistent dans l'espace discursif observé.

La projection des sous-groupes (Figure A4, reportée en annexe) visualise la cohérence interne de la typologie plutôt qu'elle ne la valide : les cinq sous-groupes étant construits à partir

d'indicateurs discursifs proches de ceux qui structurent l'ACM, leur projection dans l'espace issu des mêmes registres en cartographie la position sans en établir indépendamment la réalité. Elle fait néanmoins apparaître plusieurs éléments structurants. Le sous-groupe `standard_non` se positionne nettement du côté négatif de la première dimension, ce qui indique une faible mobilisation des registres patrimoniaux. Le sous-groupe `luxe_patrimoine` se positionne du côté positif de la première dimension et du côté inférieur de la seconde, ce qui correspond à une grammaire d'ancrage patrimonial, temporel et institutionnel. Le sous-groupe `standard_hybride` se positionne du côté positif de la première dimension et du côté supérieur de la seconde, ce qui correspond à une grammaire plus expressive et créative, combinée à une mobilisation partielle des registres d'enrichissement. Le sous-groupe `creation_auteur` occupe une position intermédiaire sur la seconde dimension, du côté positif de la première.

Le résultat le plus intéressant porte sur la position relative de `standard_hybride` et `luxe_patrimoine`. Les deux partagent le côté positif de la première dimension, c'est-à-dire le pôle enrichi de l'espace discursif. Ils s'opposent toutefois nettement sur la seconde dimension. Cette opposition aide à comprendre pourquoi les hybrides standard ne sont pas seulement du « petit luxe ». Des cas comme Guy Degrenne activent des marqueurs patrimoniaux tout en construisant un récit différent de celui des maisons de luxe patrimonial. Dans cet espace factoriel, cette catégorie hybride apparaît donc comme une voie distincte d'enrichissement : elle combine des registres d'expression identitaire, stylistique et créative avec une mobilisation partielle des registres patrimoniaux. Les deux formes partagent le pôle enrichi, mais elles mobilisent des grammaires symboliques qui ne se réduisent pas l'une à l'autre.

Pour sortir autant que possible de la circularité du codage discursif, j'ai aussi projeté des variables supplémentaires qui ne participent pas à la construction de l'espace. La projection du type d'espace commercial (Figure 2) montre que le grand magasin central se distingue du côté positif du premier axe, tandis que les centres commerciaux de couronne et généralistes, ainsi que l'espace de commerce de transit, se situent davantage du côté des faibles mobilisations patrimoniales sur le premier axe. Cette projection n'est pas circulaire : la variable type d'espace n'a servi ni à construire les sous-groupes ni à coder les registres. Elle suggère donc une correspondance entre position spatiale et position factorielle.

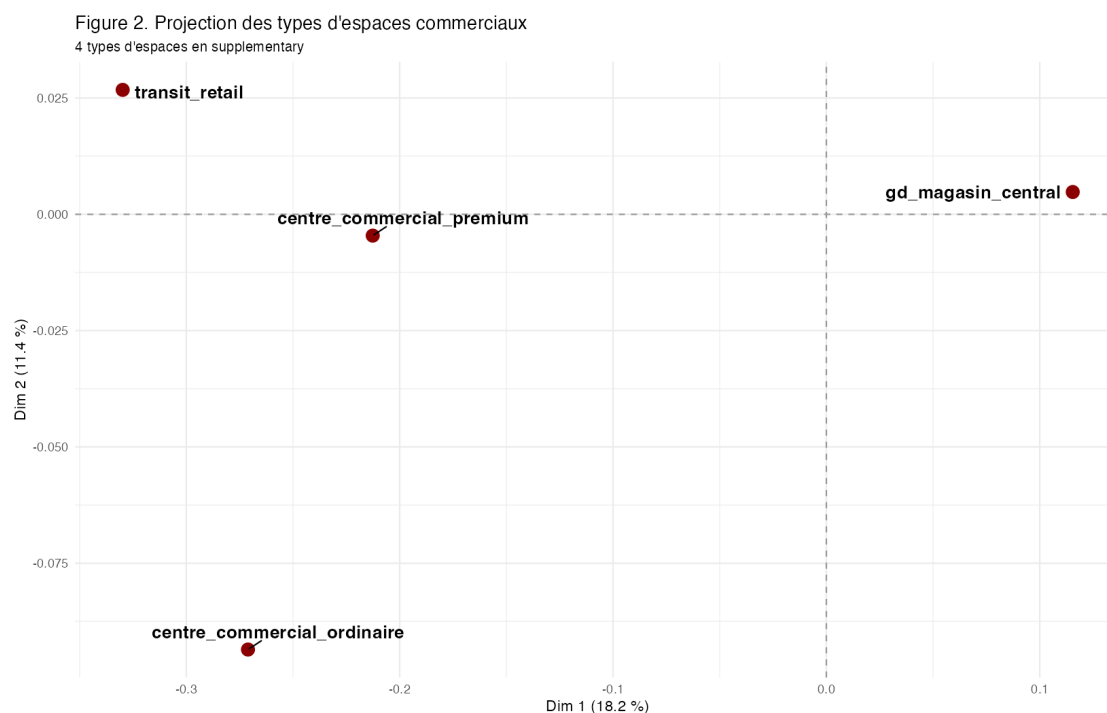


Figure 2. Projection des quatre types d'espace commercial en variables supplémentaires. Le grand magasin central se distingue du côté positif du premier axe.

La projection par centre (Figure A2, reportée en annexe) nuance la projection par type d'espace. Le Bon Marché Rive Gauche et la Samaritaine se situent du côté positif du premier axe, au plus près du pôle patrimonial et enrichi. Le BHV Rivoli, pourtant classé dans le même type, occupe une position plus centrale et plus basse sur le second axe. Cette différence rappelle que la catégorie « grand magasin central » n'est pas homogène : elle regroupe des espaces patrimoniaux fortement orientés vers le luxe, mais aussi un grand magasin plus généraliste. Le BHV Rivoli est ainsi le cas qui résiste le plus à ma typologie : la projection en confirme l'essentiel, mais elle oblige à traiter « grand magasin central » comme une catégorie composite, et non comme un type homogène.

La projection des indicateurs de prix (Figure 3) suit un gradient globalement cohérent. Les catégories de prix faibles se situent plutôt du côté des non-mobilisations discursives, tandis que les catégories de prix élevées, en particulier les plafonds de prix, se rapprochent des modalités associées au luxe explicite, au savoir-faire, à l'archive et au patrimoine. Le prix n'a pas servi à construire l'espace factoriel ; c'est pour cette raison que sa projection est informative.

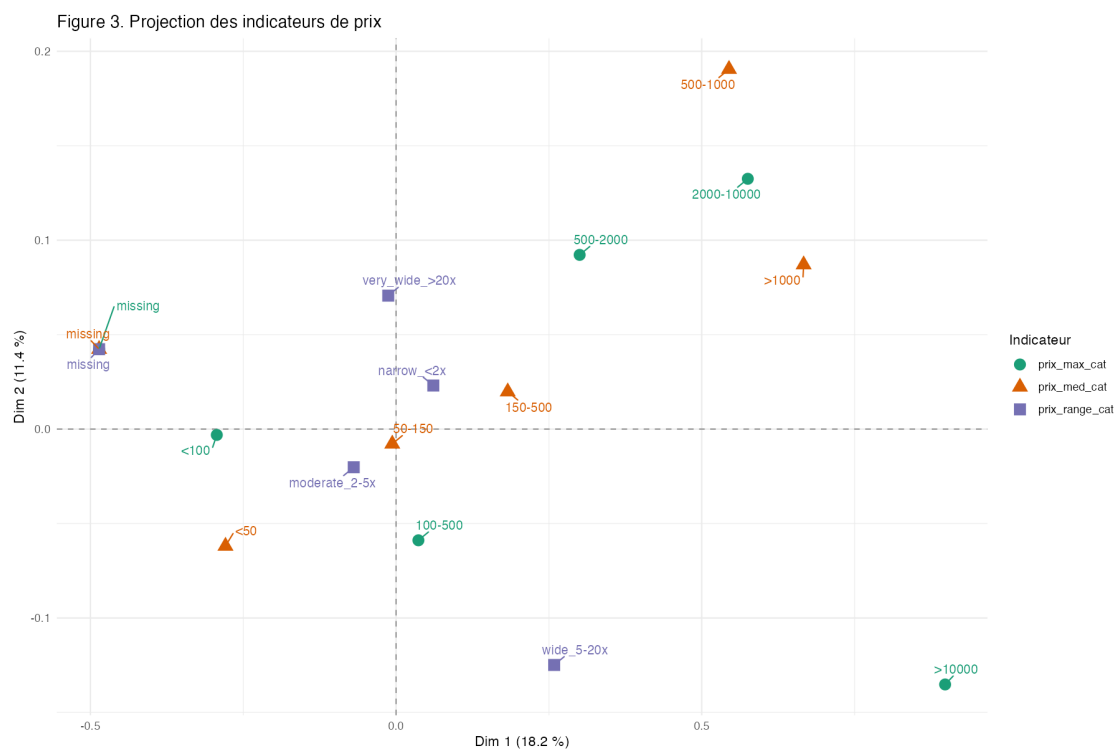


Figure 3. Projection des trois indicateurs de prix (médian, maximal, étendue) en variables supplémentaires. Le gradient des prix élevés se rapproche du pôle d'enrichissement.

La projection des cohortes de fondation (Figure A3, reportée en annexe) indique que les cohortes les plus anciennes (avant 1900, 1900-1944) se positionnent du côté du pôle d'ancrage patrimonial, tandis que les cohortes récentes (1980-1999, 2000+) se rapprochent du pôle expressif et créatif. Ce résultat est cohérent avec ce qu'on pouvait attendre : les marques anciennes semblent davantage associées aux registres patrimoniaux, les récentes aux registres expressifs et contemporains. Cette projection reste affectée par la reconstruction de l'âge : pour environ 12 % des observations, la date utilisée est celle de l'immatriculation Insee, non celle de la fondation de la marque.

3.3. Le discours et la hiérarchie des prix

Six modèles de régression linéaire (moindres carrés ordinaires) sont estimés sur les observations disposant à la fois d'un discours codé et de prix collectés. La variable dépendante change selon le modèle : le logarithme du prix médian de la marque (M1), de son prix maximal (M2, M3), du prix d'entrée de gamme (M5) ou de l'amplitude du panier observé (M6) ; M4 prend pour variable dépendante le prix relatif au secteur. Le passage au logarithme linéarise des distributions de prix fortement asymétriques et permet de lire les coefficients comme des variations relatives approximatives. Ces modèles sont utilisés comme outils de

description conditionnelle du corpus, et non comme estimations représentatives de l'ensemble du commerce francilien.

Les variables explicatives se répartissent en quatre blocs : la *forme discursive dominante* (Sous-groupe, référence `standard_non`), le *type d'espace commercial* (Espace, référence grand magasin central), le *secteur analytique* (Secteur, contrôlé mais non reporté) et la *grammaire du discours* — les quatre scores `collection_v2`, `auteur_v2`, `trend_v2`, `care_v2` et les deux covariables binaires `luxe_cov` et `origine_cov`.

M1 et M2 posent le cadre en n'introduisant que la position typologique, spatiale et sectorielle, d'abord sur le prix médian puis sur le prix maximal. M3 ajoute les variables de grammaire, pour isoler ce que la combinaison lexicale capte au-delà de la seule position typologique d'ensemble. M4 remplace le prix par le prix relatif au secteur, défini comme l'écart entre le logarithme du prix médian de la marque et le logarithme de la médiane de prix de son secteur : cette normalisation neutralise les écarts de niveau entre secteurs — un sac et un parfum n'ont pas le même prix de référence — et vérifie que les effets observés ne tiennent pas à la seule composition sectorielle du corpus. M4 reste donc sur une échelle logarithmique relative, mais son interprétation porte sur l'écart au secteur plutôt que sur le niveau absolu du prix. M5 et M6 déplacent enfin l'attention du niveau vers la forme du panier observé, en prenant pour cible le prix d'entrée puis l'amplitude max/min. Les spécifications de robustesse (sans la joaillerie, sur le prix médian, sur le 90^e percentile, avec contrôle du groupe d'entreprise, écarts-types clusterisés par marque) figurent en annexe. Je commente ici les effets les plus structurants ; les coefficients complets sont dans les tableaux.

Le tableau suivant rassemble les coefficients des quatre modèles principaux.

Variable	M1 (log prix méd.)	M2 (log prix max)	M3 (log prix max, + grammaire)	M4 (prix rel. secteur)
<code>standard_faible</code>	+0,05	-0,01	-0,04	+0,04
<code>standard_hybride</code>	+0,11 *	+0,05	0,00	+0,09
<code>creation_auteur</code>	+0,28	+0,29	+0,23	+0,24

	***	***	***	***
luxe_patrimoine	+0,32 ***	+0,35 ***	+0,23 ***	+0,25 ***
couronne (centre_commercial_premium)	-0,41 ***	-0,43 ***	-0,40 ***	-0,41 ***
généraliste (centre_commercial_ordinaire)	-0,60 ***	-0,61 ***	-0,57 ***	-0,57 ***
transit (transit_retail)	-0,58 ***	-0,64 ***	-0,61 ***	-0,58 ***
collection_v2	—	—	+0,05 ***	+0,03 **
auteur_v2	—	—	+0,09 **	+0,09 **
trend_v2	—	—	-0,01	-0,01
care_v2	—	—	-0,05 *	-0,04 *
luxe_cov	—	—	+0,32 ***	+0,25 ***
origine_cov	—	—	-0,17 **	-0,15 *

Références : sous-groupe standard_non, espace grand magasin central. Le secteur est contrôlé mais non reporté. Seuils : * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. *Champ* : observations disposant d'un discours codé et d'un indicateur de prix exploitable ; les effectifs sont de 2 350 observations pour M1 à M3 et de 2 189 pour M4, qui exclut les secteurs sans référentiel sectoriel fiable.

L'effet le plus robuste concerne la revendication explicite du luxe. La variable `luxe_cov` code la mobilisation d'un vocabulaire de luxe (haute couture, haute joaillerie, maison de luxe, fleuron du luxe). Elle reste positivement associée au plafond de prix dans le modèle principal ($\beta = +0,32$, $\exp(\beta) - 1 \approx +37\%$, $p < 0,001$). Cet effet reste stable dans les robustesses : il tient sans le secteur joaillerie, sur le prix médian, sur le prix relatif au secteur, sur le 90^e percentile,

et sous écarts-types clusterisés par marque ($p < 0,001$). C'est la variable discursive la plus fortement associée au plafond de prix, et la plus stable. Les autres associations sont plus modestes.

Deux autres scores de grammaire produisent des effets positifs plus modestes, mais nets. Le score `collection_v2`, qui agrège les marqueurs des cinq registres classiques d'enrichissement (patrimoine, ancienneté, savoir-faire, archive et, dans la version initiale du score, rareté-collection), est positivement associé au plafond de prix ($\beta = +0,05$, soit environ +5 %, $p < 0,001$). Cet effet est de faible magnitude, mais il se maintient dans l'ensemble des spécifications, y compris sous écarts-types clusterisés par marque ($p = 0,001$). Une analyse de décomposition montre qu'il est plus fort dans un modèle scores-only sans contrôle du sous-groupe ($\beta = +0,07$), ce qui indique qu'une partie de cet effet est absorbée par le sous-groupe dans la spécification complète. La collection narrative est donc associée au plafond de prix dans la direction attendue, de manière robuste mais d'ampleur limitée. Le score `auteur_v2`, qui repère la revendication d'une signature ou d'un auteur nommé, est lui aussi positivement associé au plafond de prix ($\beta = +0,09$, soit environ +9 %, $p = 0,010$), et cet effet reste significatif dans toutes les spécifications de robustesse, y compris sous écarts-types clusterisés par marque ($p = 0,030$). Le marqueur d'auteur est donc associé à un avantage tarifaire modeste mais robuste.

À l'inverse, plusieurs variables tirent le prix vers le bas ou n'y sont pas associées. Le score `care_v2`, qui agrège les marqueurs de care-émotion et de responsabilité-éthique, est négativement associé au plafond de prix ($\beta = -0,05$, $\exp(\beta)-1 \approx -5$ %, $p = 0,015$). Cet effet est négatif dans les quatre spécifications alternatives, mais sa significativité s'affaiblit sous écarts-types clusterisés par marque ($p = 0,054$) : il s'agit d'une tendance négative modérée, d'amplitude limitée. La covariable `origine_cov`, qui code la revendication d'une origine franco-parisienne, est elle aussi négativement associée au plafond de prix ($\beta = -0,17$, soit environ -16 %, $p = 0,009$), un résultat stable sous écarts-types clusterisés ($p = 0,021$). Le score `trend_v2`, enfin, n'est associé à aucun effet significatif ($p = 0,7$). La référence franco-parisienne est mobilisée par des marques très diverses, y compris accessibles ou intermédiaires, et ne suffit pas à produire un effet de prestige comparable au luxe explicite. Ces registres peuvent être lus comme des frontières symboliques actives, mais ils ne se

traduisent pas ici par un plafond tarifaire plus élevé, ce qui rejoint la distinction de Lamont et Molnár³¹ entre frontières symboliques et frontières sociales effectives.

L'effet propre du marqueur d'auteur s'ajoute à celui, plus marqué, du sous-groupe *creation_auteur* ($\beta \approx +0,23$ sur le prix maximal dans M3, $p < 0,001$). La revendication d'auteur opère ainsi à deux niveaux : comme marqueur lexical isolé, à l'effet modéré, et comme forme discursive d'ensemble, à l'effet plus fort.

Les effets de sous-groupe sont, dans l'ensemble, les plus stables. *creation_auteur* et *luxes_patrimoine* conservent un écart positif net par rapport à *standard_non* dans tous les modèles principaux (avec des coefficients comparables, de l'ordre de $+0,23$ sur le prix maximal) et sous écarts-types clusterisés par marque ($p < 0,001$). *standard_hybride* occupe une position plus ambiguë : son écart au plafond de prix devient non significatif une fois la grammaire contrôlée dans M3 ($\beta \approx 0,00$), alors qu'il reste positif et significatif sur le prix médian dans M1 ($\beta \approx +0,11$, $p = 0,029$) et sur le prix d'entrée dans M5 ($\beta \approx +0,13$, $p = 0,013$). La montée en gamme de *standard_hybride* se lit donc moins dans le plafond que dans le niveau central et le prix d'entrée de la gamme observée, mécanisme que les modèles M5 et M6 précisent.

Les effets spatiaux sont ceux qui varient le moins d'une spécification à l'autre. Par rapport au grand magasin central pris comme référence, les espaces commerciaux de couronne (*centre_commercial_premium*), les centres commerciaux généralistes (*centre_commercial_ordinaire*) et l'espace de transit (*transit_retail*) présentent tous des coefficients significativement négatifs dans les modèles principaux, avec des valeurs comprises entre $-0,40$ et $-0,61$ sur le prix maximal, et ils restent fortement significatifs sous écarts-types clusterisés par marque. Dans ces modèles, les grands magasins centraux restent associés aux prix les plus élevés, même après contrôle du secteur et des variables discursives. La position spatiale et la position discursive ne sont pas redondantes : elles capturent des dimensions partiellement indépendantes de la hiérarchie des prix. Les conclusions principales restent d'ailleurs inchangées lorsque les erreurs-types sont clusterisées par marque, ce qui prend en compte la présence d'une même enseigne dans plusieurs centres.

³¹ Lamont Michèle et Molnár Virág, « The Study of Boundaries in the Social Sciences », *Annual Review of Sociology*, 2002, vol. 28, p. 167-195.

3.4. Amplitude interne de la gamme observée : plancher, plafond et extension de gamme

Les modèles précédents portent sur des grandeurs résumées du panier de prix (prix maximal, prix médian, prix relatif au secteur). Le panier de dix produits collecté pour chaque marque ne reconstitue pas l'intégralité du catalogue ; il permet seulement d'observer la forme de la gamme retenue pour l'enquête. Trois variables prolongent l'analyse : le prix minimal observé (`prix_min_brand`), l'amplitude entre prix maximal et minimal (`prix_amplitude_brand` = max/min), et un découpage en trois niveaux d'amplitude (`ecart_prix_q3` : catalogue resserré, moyen, étendu) calé sur les quartiles de la distribution de `prix_amplitude_brand` au niveau des marques uniques. Les coefficients des modèles M5 (prix d'entrée) et M6 (amplitude) sont rassemblés dans le tableau suivant.

Variable	M5 — log prix d'entrée	M6 — log amplitude
<code>standard_faible</code>	+0,06 *	-0,10 ***
<code>standard_hybride</code>	+0,13 *	-0,13 **
<code>creation_auteur</code>	+0,24 ***	-0,01
<code>lux_e_patrimoine</code>	+0,22 ***	+0,01
<code>lux_e_cov</code>	+0,22 ***	+0,10 ***
<code>collection_v2</code>	+0,03 **	+0,02 *
<code>auteur_v2</code>	+0,07 *	+0,02
<code>care_v2</code>	-0,06 **	+0,01
<code>origine_cov</code>	-0,16 **	0,00
<code>trend_v2</code>	+0,02	-0,03

Référence des sous-groupes : `standard_non`. Le type d'espace et le secteur sont contrôlés mais non reportés ici. Seuils : * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. Les coefficients se lisent approximativement comme des variations relatives, les deux variables dépendantes étant logarithmiques. Les modèles M5 et M6 portent sur 2 319 observations.

Le modèle M5, sur le logarithme du prix d'entrée, montre que les formes les plus classiques d'enrichissement se lisent aussi dans le plancher du panier observé, et pas seulement dans son plafond. Les sous-groupes `creation_auteur` ($\beta = +0,24$, soit environ +27 %) et

`luxe_patrimoine` ($\beta = +0,22$, soit environ +25 %) ont un prix d'entrée significativement plus élevé que `standard_non`, et la revendication explicite du luxe va dans le même sens (`luxe_cov` $\beta = +0,22$, $p < 0,001$). Les scores `collection_v2` ($\beta = +0,03$, $p = 0,005$) et `auteur_v2` ($\beta = +0,07$, $p = 0,038$) relèvent eux aussi le plancher, tandis que `care_v2` ($\beta = -0,06$, $p = 0,004$) et `origine_cov` ($\beta = -0,16$, $p = 0,008$) l'abaissent. Les espaces périphériques ont, à discours et secteur comparables, un plancher nettement plus bas que le grand magasin central (entre $-0,38$ et $-0,53$). Ces formes d'enrichissement se lisent donc comme un relèvement qui touche aussi l'entrée du panier observé, et pas seulement sa pièce la plus chère.

Le modèle M6, sur le logarithme de l'amplitude (max/min), examine plus directement une hypothèse inspirée par Boltanski et Esquerre : dans le luxe étendu, la marque peut élargir sa gamme vers le bas sans renoncer à un plafond très élevé. Le résultat est nuancé. La revendication explicite du luxe est bien associée à une amplitude plus large (`luxe_cov` $\beta = +0,10$, $p < 0,001$), et `collection_v2` va dans le même sens de manière plus faible ($\beta = +0,02$, $p = 0,022$). En revanche, les sous-groupes `luxe_patrimoine` ($\beta = +0,01$, $p = 0,6$) et `creation_auteur` ($\beta = -0,01$, $p = 0,8$) ne présentent pas d'amplitude significativement plus large que `standard_non`, tandis que `standard_faible` ($\beta = -0,10$) et `standard_hybride` ($\beta = -0,13$) ont des paniers plus resserrés. L'extension de gamme apparaît donc surtout à travers la revendication lexicale directe du luxe, et non comme une propriété générale des formes patrimoniale ou auteure. Le résultat sur `standard_hybride`, qui combine un plancher plus haut (M5 : $\beta = +0,13$) avec une amplitude resserrée (M6 : $\beta = -0,13$), précise le mécanisme : la montée en gamme partielle des marques standard fonctionne, dans le panier observé, par compression vers le haut, et non par étirement des extrêmes.

Un modèle complémentaire (M7, modèle linéaire de probabilité présenté en annexe) complète cette lecture en prenant la mobilisation du registre luxe comme variable dépendante : un panier plus étendu s'accompagne d'une probabilité plus élevée de revendication du registre luxe ($\beta = +0,05$ pour l'amplitude la plus forte, $p = 0,004$), indépendamment du sous-groupe, du type d'espace et du secteur.

3.5. Position dans le secteur et profils de prix synthétiques

Les variables d'amplitude renseignent sur la forme interne du panier observé pour chaque marque. Une seconde couche de variables resitue chaque observation dans la hiérarchie de prix de son secteur. La position médiane (`position_med_sect`) découpe chaque secteur en quartiles de `prix_median_brand` et classe chaque observation en quatre niveaux, du quartile accessible au quartile haut. La position du plafond (`position_max_sect`) procède de la même manière sur le prix maximal. Les quartiles sectoriels sont calculés au niveau enseigne $\times$ centre, et non au niveau marque dédoublée. Cela revient à pondérer les marques par leur présence effective dans les centres étudiés : une enseigne transversale compte autant de fois qu'elle apparaît dans l'offre observée. Il aligne donc la métrique de référence sur la composition réelle du corpus, mais il donne plus de poids aux marques multi-implantées³².

Une variable synthétique, le profil de prix sectoriel (`profile_prix_sect`), croise ensuite la position médiane et la position du plafond en six modalités interprétatives. Les profils dits « purs » (accessible, milieu, haut) désignent les cas où la médiane et le plafond restent relativement alignés dans la hiérarchie sectorielle. Les profils dits « étirés » désignent au contraire les cas où le plafond monte beaucoup plus haut que la position médiane. Le profil `lux_e_tendu` correspond aux cas où le plafond atteint le quartile supérieur du secteur avec une amplitude interne forte ; le profil `plafond_symbolique` désigne le cas plus extrême où le prix maximal dépasse le 90^e percentile sectoriel. Ces modalités sont construites de manière exclusive : chaque observation ne relève que d'un seul profil. La modalité autre absorbe les configurations restantes. Les règles exactes de regroupement sont reportées en annexe.

La distribution de la position médiane par sous-groupe va d'abord dans le sens de la cohérence de la typologie discursive. Les observations classées en `creation_auteur` et en `lux_e_patrimoine` sont nettement plus souvent situées dans le quartile supérieur de prix médian de leur secteur, respectivement 42,9 % et 37,3 %, contre 15,2 % pour `standard_non`. Le codage qualitatif des discours ne semble donc pas renvoyer seulement à

³² Une enseigne transversale comme Levi's, présente dans huit centres du corpus, pèse donc huit fois dans le calcul des quartiles sectoriels de la mode. Ce choix aligne la métrique de référence sur la composition observée de l'offre, plutôt que sur une distribution théorique des marques considérées une à une. Il donne en revanche plus de poids aux enseignes multi-implantées.

des différences rhétoriques. Il correspond aussi à des positions différenciées dans les hiérarchies sectorielles de prix.

Sous-groupe	Q1 (accessible)	Q2	Q3	Q4 (haut)
standard_non	34,5	31,0	19,4	15,2
standard_faible	23,7	26,0	28,0	22,4
standard_hybride	23,0	24,8	26,5	25,7
creation_auteur	11,9	16,9	28,3	42,9
luxe_patrimoine	14,8	14,4	33,5	37,3
autre	35,1	27,5	21,3	16,1

Pourcentages en ligne (répartition de chaque sous-groupe entre les quartiles sectoriels de prix médian) ; secteurs « non classé » et « autre » exclus du calcul des quartiles.

Le profil synthétique précise cette relation. Le sous-groupe `luxe_patrimoine` concentre la plus forte proportion de profils luxe étendu ou plafond symbolique réunis, soit 21,1 % des cas observés. Certaines marques combinent ainsi un plafond sectoriel élevé et une forte amplitude interne du panier observé. Ce profil correspond à l'idée de luxe étendu que j'essaie de mesurer : une marque peut maintenir un plafond très élevé tout en gardant des produits d'entrée plus accessibles. Le sous-groupe `creation_auteur` présente un profil légèrement différent : il se distingue davantage par des plafonds symboliques (55 cas, soit environ 11,7 %) que par une extension ordinaire du panier observé (21 cas, soit environ 4,4 %). Cette différence suggère que la singularisation par la création peut passer moins par un étagement continu de l'offre que par quelques objets fortement valorisés, capables de fixer un seuil symbolique élevé. À l'inverse, le sous-groupe `standard_hybride` ne compte que 3,5 % d'observations en profils étirés. La montée en gamme partielle des marques standard ne se fait pas par étirement des extrêmes mais par une compression vers le haut, qui rejoint le résultat obtenu en M6.

Un dernier modèle linéaire de probabilité (M8, présenté en annexe) régresse la mobilisation du registre luxe explicite sur le profil de prix sectoriel, construit uniquement à partir des prix. Il s'agit d'un modèle descriptif : il mesure l'association conditionnelle entre le profil de prix et le discours du luxe, et ne dit rien de leur sens causal. La probabilité de mobiliser ce registre est la plus élevée pour le profil `plafond_symbolique` ($\beta = +0,17$, $p < 0,001$), à un niveau

comparable à celui des marques globalement situées en haut de leur secteur (haut_pur , $\beta = +0,16$). Le discours du luxe s'associe donc surtout aux marques qui combinent un plafond sectoriel très élevé et une forte amplitude interne, au-delà des seules marques chères en moyenne.

3.6. Marques transversales et position locale

L'analyse des marques transversales permet de revenir sur l'intuition formulée en introduction. Certaines enseignes sont présentes dans plusieurs espaces commerciaux très différenciés, mais leur position relative varie fortement selon le centre dans lequel elles apparaissent. Pour saisir cette variation, on calcule pour chaque ligne enseigne $\times$ centre un indicateur local, défini comme le logarithme décimal (base 10) du rapport entre le prix médian de la marque et la médiane des prix médians observés dans le centre. Le tableau suivant l'illustre pour trois enseignes transversales.

Enseigne	Centre	Médiane du centre	Prix médian enseigne	Position locale (log10)
Levi's	Le Bon Marché Rive Gauche	136 €	87 €	-0,20
Levi's	Italie 2	30 €	87 €	+0,46
Adidas	Le Bon Marché Rive Gauche	136 €	115 €	-0,07
Adidas	Italie 2	30 €	115 €	+0,58
Sephora	Westfield Les 4 Temps	52 €	22 €	-0,38
Sephora	Italie 2	30 €	22 €	-0,14

La médiane du centre correspond à la médiane des prix médians des enseignes observées dans le centre.

Le cas de Levi's est le plus net : la structure de prix observée reste stable, avec un prix médian d'environ 87 € et un profil sectoriel de type accessible pur, mais sa position locale varie fortement. Au Bon Marché Rive Gauche, où la médiane du centre atteint 136 €, elle reste en dessous (-0,20) ; à Italie 2, où cette médiane n'est que de 30 €, elle passe nettement au-dessus

(+0,46). L'écart de 0,66 point sur l'échelle logarithmique correspond à un rapport d'environ 4,6 entre les deux positions relatives. Il indique que la même enseigne, à structure de prix comparable, ne produit pas le même effet de position selon son environnement commercial.

Adidas se lit différemment, car la gamme observée varie elle-même selon le centre : au Bon Marché, le panier relevé présente surtout un plafond nettement plus élevé (500 € contre 130 € ailleurs), ce qui peut correspondre à des collaborations ou à des pièces signature absentes des autres relevés. La transversalité peut donc prendre deux formes : une gamme relativement stable dont seule la position locale se déplace (Levi's), ou une adaptation visible de la gamme elle-même (Adidas). Sephora, à l'inverse, relève surtout d'une diffusion dans les centres de couronne et reste partout en dessous de la médiane locale (entre -0,38 et -0,14) : elle ne traverse pas, dans le corpus étudié, les grands magasins centraux.

L'analyse transversale n'apporte pas de démonstration causale. Elle objective tout de même l'intuition formulée en introduction : l'identité de la marque ne suffit pas à fixer la position qu'elle occupe dans l'offre commerciale. La position dépend conjointement de la gamme observée et de la composition relative du centre dans lequel cette gamme prend place. Cette dépendance contextuelle justifie le choix de l'unité d'observation enseigne $\times$ centre, qui aurait été masquée si l'on avait travaillé sur des marques dédoublées. D'autres cas, non reportés dans le tableau, comme Lacoste, rappellent en outre que le classement sectoriel lui-même doit être lu avec prudence.

3.7. La sélection spatiale et l'hétérogénéité des grands magasins

L'hypothèse de sélection spatiale (H3) distingue deux mécanismes possibles : un effet de sélection des formes discursives par les types d'espace commercial, et un effet d'amplification de leur valorisation marchande. Cette hypothèse est testée par deux régressions complémentaires. Le modèle H3c, restreint aux marques standard, régresse la probabilité d'appartenir à la catégorie *standard_hybride* sur le type d'espace commercial et le secteur analytique. Le modèle H3b teste l'effet d'interaction entre le score collection et le type d'espace dans la régression sur le prix maximal.

Les résultats de H3c indiquent que, parmi les marques standard, les espaces commerciaux de couronne ont des odds nettement plus faibles d'abriter une forme *standard_hybride* que les grands magasins centraux (OR = 0,36, $p < 0,001$). Autrement dit, la forme *standard_hybride* est significativement plus concentrée dans les grands magasins

centraux que dans les espaces de couronne. Les comparaisons avec les autres types d'espace ne sont pas significatives : *centre_commercial_ordinaire* (OR = 0,34, p = 0,14) et *transit_retail* (OR = 0,69, p = 0,4), ce qui doit être lu avec prudence compte tenu de la faible taille de ces catégories. Du côté sectoriel, la mode présente un odds-ratio significativement supérieur (OR = 2,44, p = 0,005). La forme *standard_hybride* se concentre ainsi préférentiellement dans le grand magasin central, et préférentiellement dans la mode.

Les résultats de H3b sur l'interaction collection × espace ne soutiennent pas l'hypothèse d'amplification. Sur les trois coefficients d'interaction, deux ne se distinguent pas du grand magasin central (centre commercial généraliste et espace de transit, non significatifs) ; le troisième, pour l'espace commercial de couronne, est faiblement significatif et négatif ($\beta = -0,05$, p = 0,040). Aucun n'est positif. Autrement dit, aucun espace n'amplifie l'association entre la grammaire collectionnelle et le prix, l'espace de couronne tendant même à l'atténuer légèrement. Ce résultat est important pour l'interprétation : il indique que l'espace commercial ne modifie pas vraiment l'association entre récit collectionnel et prix. L'espace ne rend donc pas un même récit patrimonial plus cher ; il sélectionne plutôt les marques qui le portent déjà. Dans les modèles, le grand magasin central ne valorise pas davantage un récit donné qu'un centre commercial périphérique ; il y concentre surtout, plus densément, les marques qui le produisent.

Le cas qui gêne le plus la typologie est le BHV Rivoli. Traité jusqu'ici comme un grand magasin central homogène avec Le Bon Marché et la Samaritaine, il s'en écarte nettement à l'examen. Au Bon Marché Rive Gauche, le sous-groupe *standard_hybride* représente 65 observations, soit 5,9 % du centre. À la Samaritaine, il représente 29 observations, soit 5,6 % du centre. Ces proportions contrastent fortement avec celles observées au BHV Rivoli, qui ne comptabilise que 4 observations *standard_hybride*, soit 1,0 % du centre, et dont le sous-groupe autre atteint 44,6 %, en raison notamment d'activités moins directement comparables aux marques de mode ou de luxe analysées ici (maison, bricolage, loisirs). Les autres centres présentent des proportions encore plus faibles : 2,0 % à Italie 2, 1,4 % à Westfield Les 4 Temps. La distribution suggère que la catégorie *grand_magasin_central* regroupe deux sous-types fonctionnellement distincts plutôt qu'un ensemble homogène. Le Bon Marché et la Samaritaine fonctionnent comme grands magasins de curation luxe, mode et auteur, où la concentration discursive du standard hybride est élevée. Le BHV Rivoli fonctionne selon une logique plus généraliste, ce qui suggère une part importante de

catégories moins directement prises dans la grammaire d'enrichissement étudiée ici (maison, bricolage, loisirs), où l'hybridation standard est peu présente.

Cette différenciation interne invite à préciser la lecture hétérotopique des grands magasins : Le Bon Marché et la Samaritaine se rapprochent davantage du pôle touristique-luxueux, où l'inscription patrimoniale du bâtiment et la curation de l'assortiment convergent, tandis que le BHV Rivoli répond à une logique plus généraliste. Ce décalage signale moins une erreur de classement qu'une limite utile de la typologie : il invite à distinguer au moins deux régimes internes de grand magasin central, qu'il faudra raffiner dans de futurs travaux.

3.8. Ce que la base permet de retenir

À ce stade, je retiens surtout trois résultats.

Le premier résultat est celui de la classification typologique. La typologie discursive construite dans le mémoire trouve un appui empirique dans les modèles de prix : les observations inscrites dans une forme dominante d'enrichissement, qu'elle relève du patrimoine ou de la création d'auteur, présentent des plafonds de prix nettement plus élevés que les marques standard sans enrichissement. Les sous-groupes *luxe_patrimoine* et *creation_auteur* conservent un écart positif net par rapport à *standard_non* dans tous les modèles principaux et dans toutes les robustesses. C'est le résultat le plus solide du mémoire au niveau typologique.

Le deuxième résultat concerne la différenciation interne de la grammaire de l'enrichissement. L'ACM fait apparaître deux dimensions distinctes, soit deux manières de mettre en valeur : l'enrichissement patrimonial et artisanal classique d'un côté, une grammaire plus expressive et créative de l'autre, où se regroupent notamment l'identité, le style de vie, la signature, mais aussi le care. Ces deux grammaires n'entretiennent pas les mêmes associations avec les prix. La revendication explicite du luxe est nettement et stablement associée à des prix plus élevés ; le marqueur d'auteur et la collection narrative le sont aussi, mais de manière plus modeste. Le care, la responsabilité et l'origine franco-parisienne sont au contraire associés à des prix plus faibles. Les modèles sur le prix d'entrée et l'amplitude précisent ce tableau : les formes classiques d'enrichissement se lisent aussi dans le prix d'entrée du panier observé, et la rhétorique luxueuse explicite s'accompagne d'une amplitude plus large, cohérente avec l'hypothèse du luxe étendu retenue ici. Plusieurs grammaires de valorisation coexistent donc,

et toutes ne se traduisent pas en prix plus élevés. Le care en est l'exemple le plus visible, avec un effet d'ampleur modérée.

Le troisième résultat est celui de la sélection spatiale. Les formes *standard_hybride* sont plus concentrées dans les grands magasins centraux que dans les espaces commerciaux de couronne, mais la différenciation se fait par sélection et non par amplification : un récit patrimonial donné n'est pas mieux valorisé tarifairement parce qu'il se trouve en grand magasin central. Les résultats ne montrent donc pas une transformation de la pente économique du discours par l'espace ; ils vont plutôt dans le sens d'une sélection en amont. Delphine Dion et Stéphane Borraz³³ ont décrit, dans les magasins de luxe, comment le servicescape façonne les subjectivités de classe ; mes données ne permettent pas une transposition directe, puisqu'elles portent sur la composition de l'assortiment et non sur l'expérience consommatrice. Elles permettent plutôt de déplacer cette intuition au niveau de l'offre : dans mes données, le grand magasin central parisien apparaît comme un espace où se concentrent davantage les formes discursives enrichies.

Pris ensemble, les résultats empêchent de réduire le prix observé à une simple opposition luxe / non-luxe. La forme discursive compte, sans suffire : certains registres patrimoniaux et le luxe explicite vont avec des plafonds élevés, tandis que le care ou l'origine franco-parisienne tirent plutôt vers le bas, et les types d'espace agissent surtout par la composition des enseignes qu'ils accueillent.

Conclusion

Ce mémoire est parti d'un problème simple : les outils statistiques disponibles décrivent assez bien les activités commerciales, beaucoup moins la position sociale relative des enseignes dans un assortiment. Pour traiter ce décalage, j'ai construit une base enseigne × centre sur neuf espaces commerciaux franciliens, composée de 2 978 observations pour 2 173 marques uniques. L'analyse met en relation trois dimensions rarement observées ensemble : les discours institutionnels des marques, leurs prix publics et leur position dans des espaces commerciaux différenciés.

³³ Dion Delphine et Borraz Stéphane, « Managing Status: How Luxury Brands Shape Class Subjectivities in the Service Encounter », *Journal of Marketing*, 2017, vol. 81, n° 5, p. 67-85.

Le résultat principal n'est pas que les marques riches parlent davantage de patrimoine. Il est plus circonscrit : dans ce corpus, certaines formes de discours institutionnel, surtout le luxe explicite et les formes patrimoniales fortes, sont associées à des plafonds de prix plus élevés, tandis que d'autres registres comme le care ou l'origine franco-parisienne ne le sont pas, ou le sont négativement. Quant aux types d'espace, ils semblent surtout sélectionner les formes discursives présentes plutôt que transformer leur rendement tarifaire. Ce contraste tient d'ailleurs surtout à l'opposition entre Le Bon Marché et la Samaritaine, d'un côté, et les espaces plus périphériques de l'autre ; Westfield Rosny 2 et Italie 2 restent trop peu fournis en marques enrichies pour autoriser une généralisation forte. Le Bon Marché, la Samaritaine et le BHV Rivoli ne relèvent pas du même régime commercial : Le Bon Marché et la Samaritaine concentrent la forme standard_hybride bien plus que le BHV Rivoli, et les traiter comme une seule catégorie masque une opposition interne décisive. Les résultats valident l'intérêt descriptif des variables construites : elles isolent des écarts de prix, des différences de grammaire discursive et des effets de sélection spatiale que les nomenclatures d'activité ne captent pas.

Ces résultats ont un périmètre assez précis, et trois limites en fixent la portée, par ordre d'importance décroissante. La première tient à la structure de l'échantillon : les grands magasins centraux y pèsent trop lourd, et les inférences sur les centres ordinaires reposent sur des effectifs réduits dont la puissance statistique reste faible. La deuxième tient aux prix : leur visibilité publique varie d'une marque à l'autre, ce qui fragilise les comparaisons fines de gamme. La troisième tient au codage : il s'appuie sur un classifieur supervisé entraîné sur annotations manuelles documentées, sans vérification par accord inter-codeurs. La typologie spatiale, enfin, reste une construction de recherche assumée, qui ne reprend aucune nomenclature professionnelle. Ces limites fixent la portée du mémoire : il décrit un corpus situé, sans prétendre produire une typologie générale du commerce francilien.

Ce travail laisse surtout trois pistes ouvertes. La première serait comparative ou longitudinale : étendre l'enquête à d'autres métropoles, ou la suivre dans le temps, pour tester la portée de la grammaire de l'enrichissement au-delà du cas parisien. La deuxième consisterait à mieux analyser les dispositifs de jugement qui rendent les prix acceptables : réputation, rareté, mise en comparaison, attentes autour des marques. Beckert et Karpik offrirait ici des appuis

utiles, mais ce travail dépasse le cadre de ce mémoire.³⁴³⁵La troisième, la plus directement ouverte par mes données, est la différenciation interne des grands magasins centraux : le contraste entre Le Bon Marché et la Samaritaine d'un côté, le BHV Rivoli de l'autre, suggère plusieurs régimes commerciaux internes, que la notion d'hétérotopie pourrait aider à analyser plus finement.

Ce mémoire aura surtout montré une chose : l'offre commerciale peut être décrite comme un espace social dès lors qu'on cesse de la traiter comme une simple liste de magasins. Les discours de marque, les prix et les types d'espace ne se recouvrent pas entièrement ; c'est justement dans leurs écarts que l'assortiment devient sociologiquement lisible. Il ne propose qu'une opérationnalisation limitée du cadre de l'enrichissement de Boltanski et Esquerre, qui demanderait à être éprouvée sur d'autres terrains que le seul cas francilien.

Bibliographie

Références scientifiques

1. Beckert Jens, *Imagined Futures: Fictional Expectations and Capitalist Dynamics*, Cambridge (MA), Harvard University Press, 2016.
<http://www.jstor.org/stable/j.ctvjnrvrw>.
2. Benjamin Walter, *Paris, capitale du XIXe siècle. Le livre des passages*, Paris, Éditions du Cerf, 1989.
3. Boltanski Luc et Esquerre Arnaud, *Enrichissement. Une critique de la marchandise*, Paris, Gallimard, 2017.
4. Bourdieu Pierre, *La Distinction. Critique sociale du jugement*, Paris, Éditions de Minuit, 1979.
5. Bourdieu Pierre, *Les structures sociales de l'économie*, Paris, Seuil, 2000.

³⁴ Beckert Jens, *Imagined Futures: Fictional Expectations and Capitalist Dynamics*, Cambridge (MA), Harvard University Press, 2016, p. 61-75 et p. 89-93.

³⁵ Karpik Lucien, « Éléments de l'économie des singularités » dans *Traité de sociologie économique*, Paris, Presses universitaires de France, 2013, p. 169-177.

6. Bourdieu Pierre et Delsaut Yvette, « Le couturier et sa griffe : contribution à une théorie de la magie », *Actes de la recherche en sciences sociales*, 1975, vol. 1, n° 1, p. 7-36, DOI : 10.3406/arss.1975.2447.
7. Chabault Vincent, *Sociologie du commerce*, Paris, Armand Colin, 2024.
8. Coëffé Vincent et Morice Jean-René, « The Parisian department store as a paradigmatic place for interactions between tourism and shopping: the production of a heterotopia », *Belgeo*, 2020, n° 1, DOI : 10.4000/belgeo.43367.
9. Desrosières Alain, *La Politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte, 1993.
10. Dion Delphine et Borraz Stéphane, « Managing Status: How Luxury Brands Shape Class Subjectivities in the Service Encounter », *Journal of Marketing*, 2017, vol. 81, n° 5, p. 67-85, DOI : 10.1509/jm.15.0291.
11. Fleury Antoine, Mathian Hélène et Saint-Julien Thérèse, « Définir les centralités commerciales au cœur d'une grande métropole : le cas de Paris intra-muros », *Cybergeo*, 2012, n° 588, DOI : 10.4000/cybergeo.25107.
12. Foucault Michel, « Des espaces autres », *Empan*, 2004, vol. 54, n° 2, p. 12-19, DOI : 10.3917/empa.054.0012.
13. Goss Jon, « The “Magic of the Mall”: An Analysis of Form, Function, and Meaning in the Contemporary Retail Built Environment », *Annals of the Association of American Geographers*, 1993, vol. 83, n° 1, p. 18-47, DOI : 10.1111/j.1467-8306.1993.tb01921.x.
14. Karpik Lucien, « Éléments de l'économie des singularités » dans *Traité de sociologie économique*, Paris, Presses universitaires de France, 2013, p. 163-206, DOI : 10.3917/puf.stein.2013.01.0163.
15. Lahire Bernard, *La Culture des individus. Dissonances culturelles et distinction de soi*, Paris, La Découverte, 2006.
16. Lamont Michèle et Molnár Virág, « The Study of Boundaries in the Social Sciences », *Annual Review of Sociology*, 2002, vol. 28, p. 167-195, DOI : 10.1146/annurev.soc.28.110601.141107.

17. Nelson Laura K., « Computational Grounded Theory: A Methodological Framework », *Sociological Methods & Research*, 2020, vol. 49, n° 1, p. 3-42, DOI : 10.1177/0049124117729703.
18. Roueff Olivier, « Les homologues structurales : une magie sociale sans magiciens ? La place des intermédiaires dans la fabrique des valeurs » dans *Trente ans après La Distinction*, de Pierre Bourdieu, Paris, La Découverte, 2013, p. 153-164.
19. Thévenot Laurent, « Les investissements de forme », dans Thévenot Laurent (dir.), *Conventions économiques*, Paris, Presses universitaires de France, 1986, p. 21-71.
20. Zukin Sharon, *Landscapes of Power. From Detroit to Disney World*, Berkeley, University of California Press, 1991.

Autres références

1. Bédoué Matthieu et Cohen Clément, « Le commerce de proximité : des pôles plus florissants en périphérie qu'en centre-ville », *Insee Première*, 2021, n° 1858, <https://www.insee.fr/fr/statistiques/5386414>, consulté le 3 juin 2026.
2. Cazaubiel Arthur et Guymarc Gaël, « La déprise du commerce de proximité dans les centres-villes des villes de taille intermédiaire », *Insee Première*, 2019, n° 1782, <https://www.insee.fr/fr/statistiques/4248184>, consulté le 3 juin 2026.

Sources et bases de données

1. ActiveTigger, outil d'annotation et de classification assistée de textes en sciences sociales computationnelles, développé par l'équipe ActiveTigger autour de Julien Boelaert et Étienne Ollion, documentation en ligne : <https://activetigger.com/documentation/>, consulté le 3 juin 2026.
2. Base SIRENE de l'Insee, accessible via l'API publique recherche-entreprises.api.gouv.fr.
3. Sites institutionnels des marques présentes dans les neuf espaces commerciaux étudiés.

4. Annuaires en ligne des centres commerciaux Le Bon Marché Rive Gauche, La Samaritaine, BHV Rivoli, BHV Parly 2, Westfield Les 4 Temps, Westfield Vélizy 2, Westfield Rosny 2, Westfield Aéroville, Italie 2.

Annexes

Les annexes au mémoire comprennent :

- Annexe A : Exemples de codage du discours (forme dominante, intensité, registres lexicaux).
- Annexe B : Figures supplémentaires de l'analyse des correspondances multiples.
- Annexe C : Modèles linéaires de probabilité (M7 et M8).
- Annexe D : Distribution détaillée du corpus (type d'espace, secteur, cohorte de fondation).
- Annexe E : Tableau récapitulatif des variables construites.
- Le codebook complet, les scripts d'analyse (R et Python), la base de données consolidée et les tableaux de robustesse détaillés des régressions, trop volumineux ou de format différent, sont fournis dans une pièce jointe séparée intitulée « LI_Zhuoran_annexe ».

Annexe A. Exemples de codage du discours

Le codage du discours articule trois schèmes : la forme dominante (logique d'ensemble du discours), l'intensité d'enrichissement (degré de mobilisation) et les registres lexicaux (contenu précis détecté par dictionnaires). Trois marques illustrent cette articulation ; les marqueurs détectés figurent entre parenthèses.

Loewe (maroquinerie de luxe). Extrait : « Fondée en 1846 par un groupe d'artisans espagnols à Madrid, LOEWE est née du monde de l'artisanat et du cuir. Aujourd'hui, cet héritage demeure au cœur de la Maison et inspire la création des collections. » Marqueurs : ancienneté (« Fondée en 1846 »), savoir-faire (« artisans », « artisanat », « cuir »), patrimoine (« héritage », « Maison »). Forme dominante : `luxe_patrimoine` — l'enrichissement patrimonial et artisanal organise l'ensemble du discours.

Gerard Darel (prêt-à-porter). Extrait : « Fondée en 1971, Gerard Darel ouvre sa première boutique à Saint-Germain-des-Prés, où les Parisiens se mêlent aux artistes et aux créateurs avant-gardistes. Paris est la source première d'inspiration des designers. » Marqueurs : auteur-

design (« créateurs », « designers »), ancienneté (« Fondée en 1971 »), origine franco-parisienne (« Paris »). Forme dominante : création_auteur.

Guy Degrenne (arts de la table). Extrait : « Maison française familiale depuis 1948, Guy Degrenne sublime l'art de recevoir en conjuguant héritage et innovation. Au cœur de nos Manufactures, le savoir-faire artisanal rencontre la précision industrielle pour donner vie à des pièces d'exception. » Marqueurs : patrimoine (« héritage »), ancienneté (« depuis 1948 »), savoir-faire (« savoir-faire artisanal »), origine franco-parisienne (« française »). Forme dominante : standard, mais intensité d'enrichissement forte, d'où le sous-groupe standard_hybride.

Annexe B. Figures supplémentaires de l'analyse des correspondances multiples

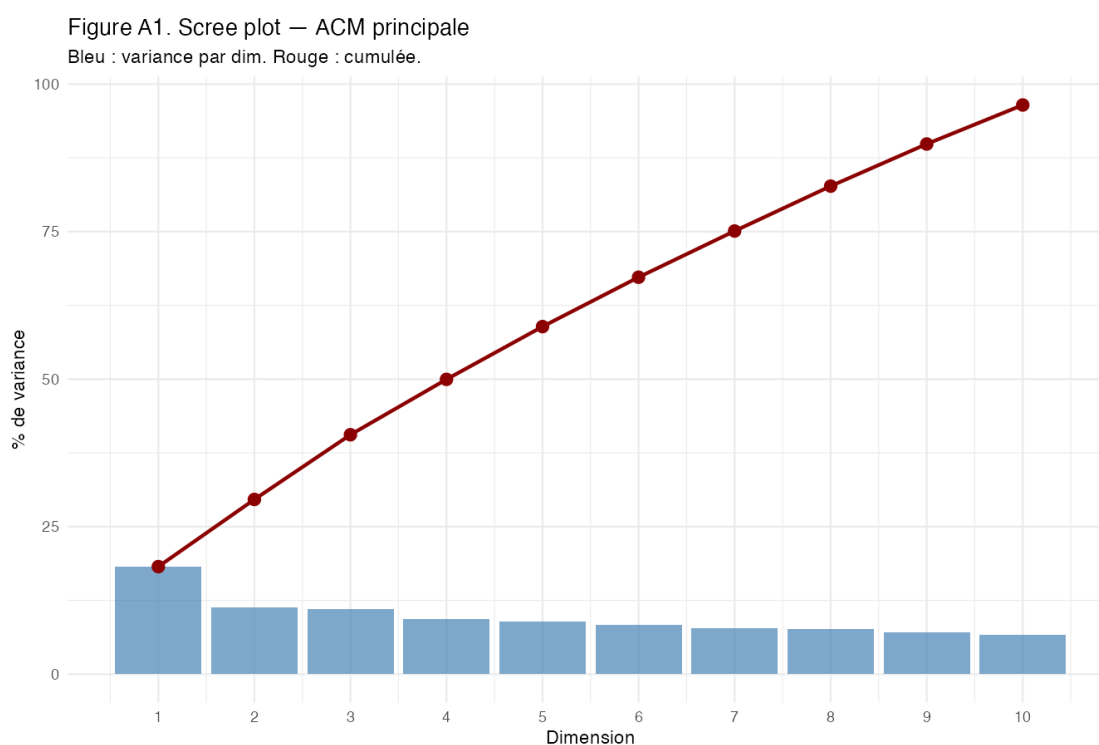


Figure A1. Scree plot de l'ACM principale. Les barres indiquent la variance expliquée par dimension ; la courbe, la variance cumulée.

Figure A2. Projection des 9 centres commerciaux
Position de chaque centre dans l'espace factoriel ; révèle l'hétérogénéité interne du groupe
gd_magasin_central.

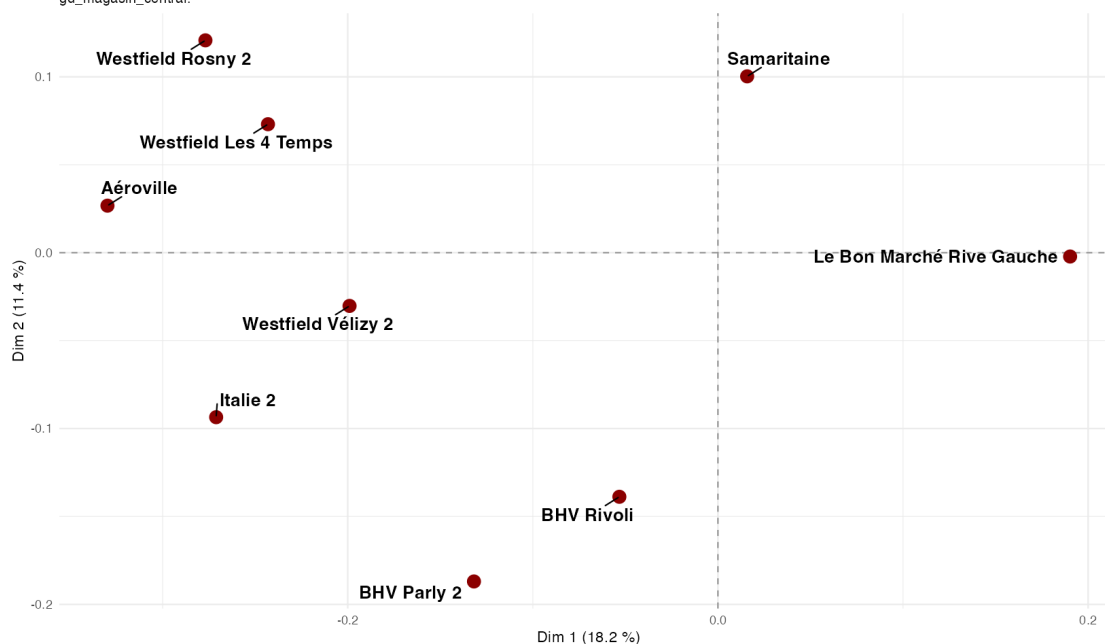


Figure A2. Projection des neuf centres en variables supplémentaires dans l'espace factoriel.
Lecture complémentaire de la projection par type d'espace (Figure 2).

Figure A3. Projection des cohortes de fondation
Variable supplémentaire — cohorte issue d'une cascade de sources

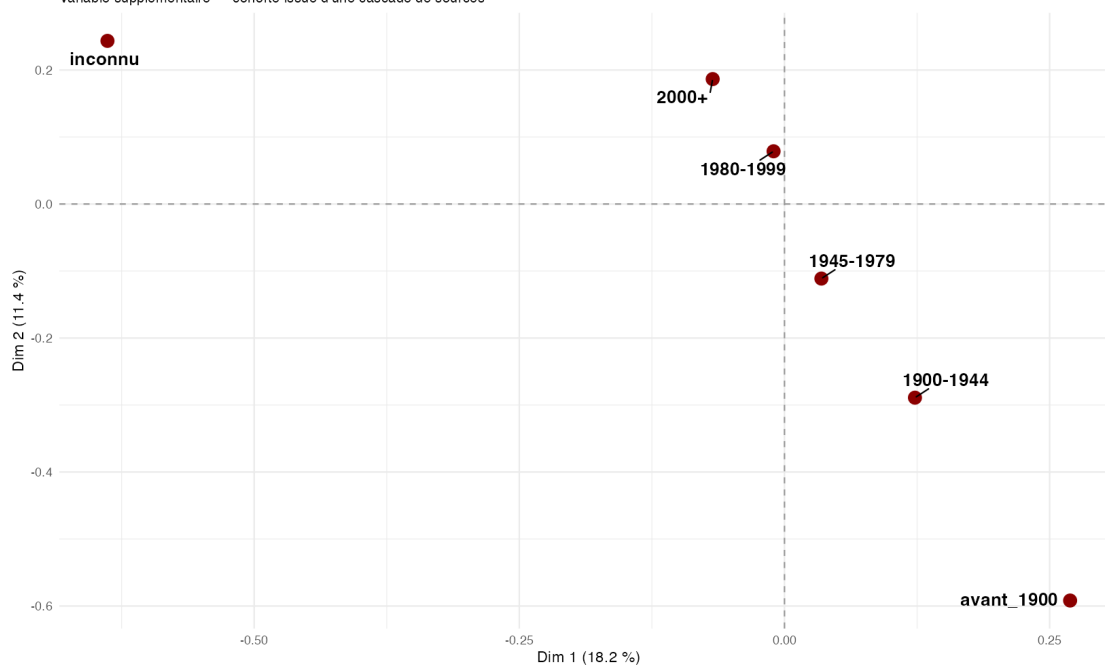


Figure A3. Projection des six cohortes de fondation en variables supplémentaires. Les cohortes anciennes se rapprochent du pôle d'ancrage patrimonial, les récentes du pôle expressif et créatif.

Figure A4. Position des formes discursives dans l'espace de l'enrichissement

Projection illustrative. Les sous-groupes sont des variables supplémentaires ; leur position situe la typologie sans constituer une validation indépendante.

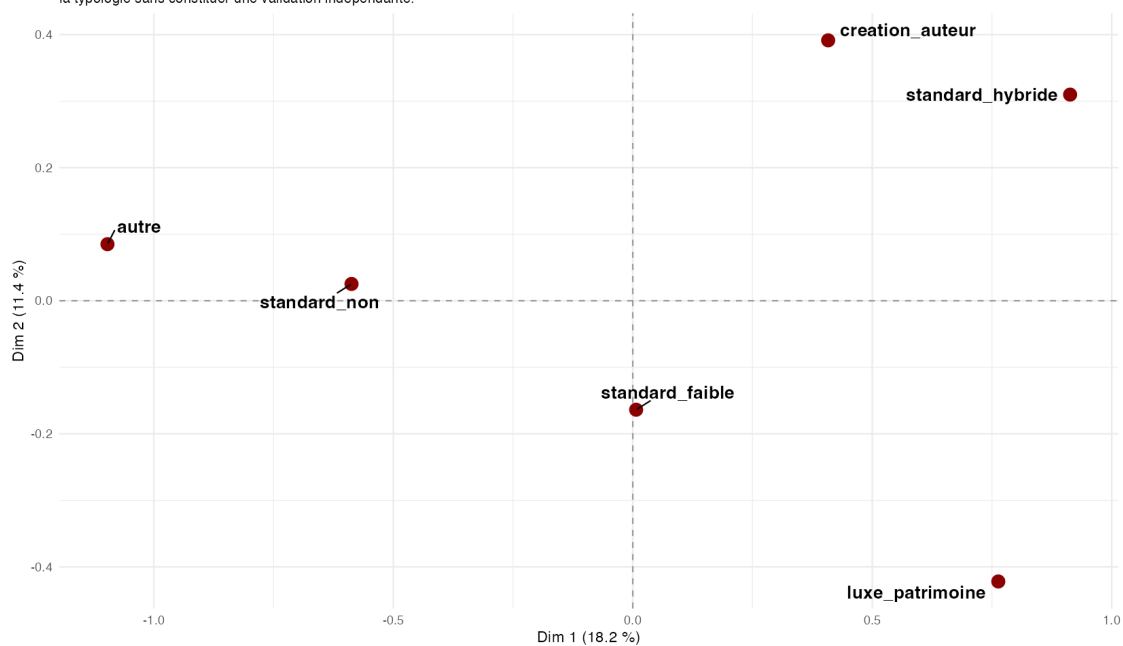


Figure A4. Projection des cinq sous-groupes en variables supplémentaires dans l'espace factoriel. La position visualise la cohérence interne de la typologie sans constituer une validation indépendante de cette typologie.

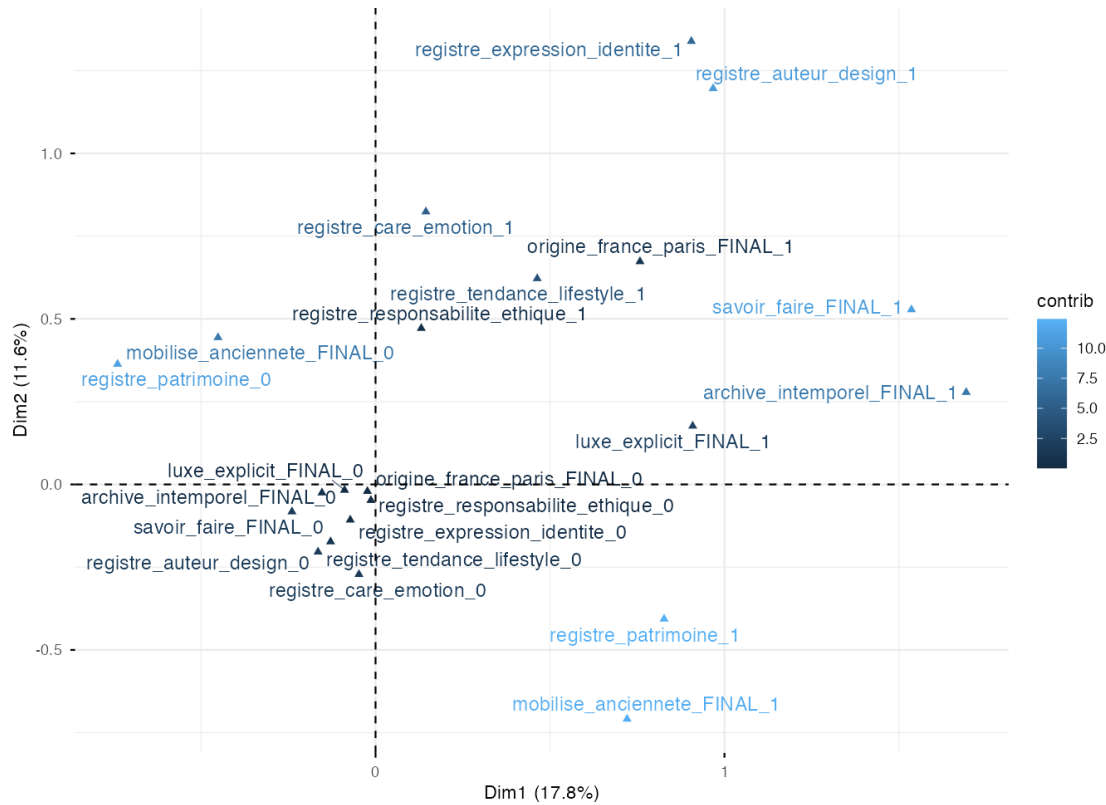


Figure A5. Analyse des correspondances multiples des onze registres, recalculée sur les marques dédoublées (une ligne par marque, $N = 1\,875$). Inerties de 17,8 % et 11,6 % pour les deux premières dimensions ; coordonnées de la première dimension quasiment inchangées par rapport à la Figure 1 (corrélation $\approx 0,98$). La structure de l'espace discursif ne dépend donc pas de la multiplicité des observations enseigne $\times$ centre.

Annexe C. Modèles linéaires de probabilité (M7 et M8)

Le modèle M7 prend la présence du registre luxe explicite (`luxe_cov`, codée 0/1) comme variable dépendante et `ecart_prix_q3` comme prédicteur principal, à sous-groupe, espace et secteur égaux. Par rapport au niveau d'amplitude le plus bas, le niveau moyen est associé à une probabilité un peu plus élevée de mobiliser le registre luxe ($\beta = +0,03$, $p = 0,032$), et le niveau élevé à une probabilité plus élevée encore ($\beta = +0,05$, $p = 0,004$). Un panier plus étendu s'accompagne donc d'une probabilité plus élevée de revendication du registre luxe, indépendamment de la classification typologique et du secteur.

Le modèle M8 régresse la présence du registre luxe explicite sur le profil de prix sectoriel, à sous-groupe, espace et secteur égaux ; ces profils sont construits uniquement à partir des prix. Par rapport au profil accessible pur, le profil milieu pur n'est pas significatif ($\beta = +0,03$, $p = 0,10$) ; le profil haut pur est positif et net ($\beta = +0,16$, $p < 0,001$) ; le profil luxe étendu est

intermédiaire ($\beta = +0,10$, $p < 0,001$) ; le profil plafond symbolique présente le coefficient le plus élevé ($\beta = +0,17$, $p < 0,001$). La probabilité de mobiliser le registre luxe est donc au moins aussi élevée lorsque la marque rend visible un plafond sectoriel très élevé combiné à une forte amplitude interne.

Profil de prix sectoriel	β (probabilité de mobiliser <i>luxe_cov</i>)
accessible_pur (réf.)	—
milieu_pur	+0,03 (ns)
haut_pur	+0,16 ***
luxe_etendu	+0,10 ***
plafond_symbolique	+0,17 ***
autre	+0,07 ***

Modèle linéaire de probabilité. Contrôles : sous-groupe, type d'espace, secteur (non reportés).
Seuils : *** $p < 0,001$; (ns) non significatif.

Annexe D. Distribution détaillée du corpus (type d'espace, secteur, cohorte de fondation)

Catégorie	N	%
<i>Type d'espace</i>		
Grand magasin central	2 026	68,0
Espace commercial de couronne	724	24,3
Espace de commerce de transit	130	4,4
Centre commercial généraliste	98	3,3
<i>Secteur</i>		
Mode	602	20,2
Services spécialisés	454	15,2
Beauté	421	14,1
Maroquinerie-chaussure	319	10,7
Joannerie-horlogerie	244	8,2
Autre	214	7,2

Bricolage	201	6,7
Enfant	196	6,6
Maison	181	6,1
Lingerie	138	4,6
Non classé	8	0,3
<i>Cohorte de fondation</i>		
Avant 1900	296	9,9
1900–1944	262	8,8
1945–1979	476	16,0
1980–1999	602	20,2
2000 et après	1 270	42,6
Inconnue	72	2,4

Annexe E. Tableau récapitulatif des variables construites

Variable	Source principale	Fonction analytique	Limite principale
Forme discursive dominante	Sites institutionnels des marques	Capter l'architecture générale du discours institutionnel	Codage interprétatif
Intensité d'enrichissement	Sites institutionnels des marques, à partir de la forme dominante et des marqueurs lexicaux	Mesurer le degré de mobilisation d'une grammaire d'enrichissement	Variable construite en partie à partir d'autres codages
Sept registres lexicaux	Sites institutionnels, dictionnaires regex	Repérer les ressources symboliques mobilisées	Sensible aux formulations atypiques
Registres lexicaux resserrés	Sites institutionnels, regex plus stricts	Tester une version plus prudente du codage lexical	Risque de sous-détection
Prix médian et	Catalogues officiels	Mesurer le niveau de prix	Comparabilité

maximal par marque	des marques	et le plafond de valorisation	limitée entre catégories de produits
Position enseigne × centre	Annuaire des centres commerciaux	Observer la présence d'une même enseigne dans différents assortiments	Présence binaire : pas la surface, pas le chiffre d'affaires
Type d'espace commercial	Typologie construite dans l'enquête	Distinguer grand magasin central, espace commercial de couronne, centre généraliste, commerce de transit	Typologie construite, donc discutable
Secteur analytique	Catégorie du centre, NAF, catégorie de produit	Contrôler les effets sectoriels	Regroupement imparfait
Cohorte de fondation	Fondation objective, fondation revendiquée, date Insee	Tester l'effet d'ancienneté	Date administrative et fondation réelle ne coïncident pas toujours
Appartenance à un groupe	Recherche manuelle, sources administratives et sources d'entreprise	Contrôler l'effet des grands groupes	Couverture partielle