

# Et l'Homme... Réprima la Femme

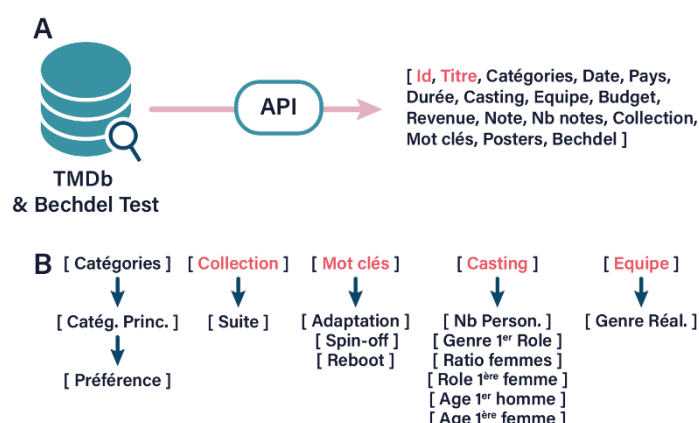
« *Un film Ghostbusters qu'avec des femmes ! Mais que se passe-t-il ??* »

Donald Trump

Ces dernières années, beaucoup de voix sur les réseaux sociaux s'élèvent pour critiquer une « surreprésentation » des femmes et/ou des personnes racisées dans les films. Ces critiques pourraient malheureusement influencer la politique des plateformes de *streaming* qui de plus en plus dictent le futur du cinéma. L'objectif de ce projet était alors d'explorer la base de données du site The Movie Database, afin d'observer s'il existe une influence de la représentation féminine sur le succès critique ou la rentabilité d'un film. Pour répondre à cette question j'ai réalisé un *clustering* non-supervisé par la méthode des *k-means* permettant de regrouper de manière automatique les différents films en fonction de nombreuses variables objectives, telles que le budget, la catégorie, le genre du premier rôle, le nombre de personnage, etc. L'étude de différents clusters m'a alors permis de mettre en évidence que les hommes sont plutôt conservateurs et contestent l'augmentation de la représentation féminine dans des rôles habituellement donnés aux acteurs.

Si vous demandiez à des passant·e·s dans la rue de définir un film féminin, beaucoup vous répondraient une « Comédie-romantique », une « Comédie-Musicale », tandis que d'autres vous citeraient des films tels que « Twilight » ou pourquoi pas « Casablanca ». Pours stéréotypes, vous penseriez ? Une étude réalisée par une équipe allemande (**Wühr, et al, 2017**) tend pourtant à confirmer qu'il existe une préférence pour certaines catégories de films liée au genre des spectateurs. Tandis que les films d'actions, d'horreur ou de science-fiction sont préférés par les hommes ; les films romantiques ou les drames attirent plutôt les femmes. Nous pourrions alors nous demander pourquoi de telles préférences ? Est-ce lié à l'attrait de la testostérone masculine pour l'action ? Ou simplement car les hommes aiment s'imaginer être Keanu Reeves dans *John Wick* ? Le « male gaze » est un concept qui désigne l'empreinte laissée par la vision masculine sur la culture visuelle dominante, notamment le cinéma. Les réalisateurs, majoritairement masculins, véhiculent alors les stéréotypes de genre encore fortement ancrés dans nos sociétés modernes. Les femmes sont alors des héroïnes dramatiques, des êtres sensibles cherchant l'amour d'un homme ou bien des femmes fatales définies par leurs corps, érigé en trophée par le héros. Cette vision d'un archétype féminin offerte par les réalisateurs justifie la surreprésentation des hommes au cinéma, et influence nos habitudes de consommation. Un cinéma fait par des hommes pour les hommes ?

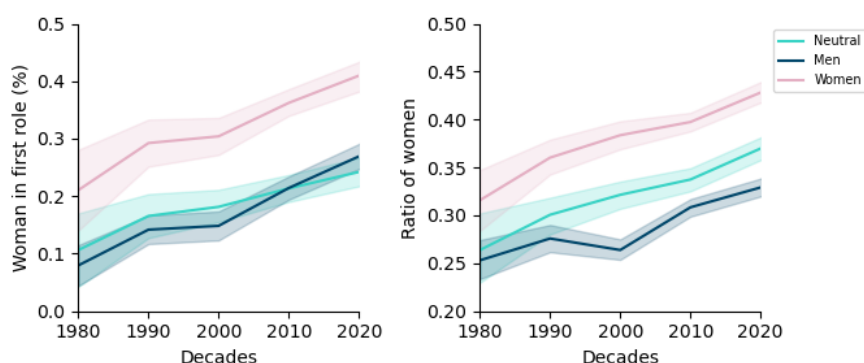
D'après l'association américaine du cinéma, 52% des spectateurs-rices sont pourtant des femmes (**MPAA, 2016**) qui se voient alors imposer cette perception masculine, au risque d'ancrer encore plus profondément cet archétype. Ces dernières années, de nombreux problèmes sociétaux furent soulevés par différents mouvements. Le scandale de l'affaire Weinstein et l'explosion du mouvement « Me Too » en 2017, ou alors les marches « Black Lives Matter » à partir de 2013, auront bousculé notre société et incité les entreprises à réfléchir à ces inégalités. Par simple « capitalisme éveillé » ou réelle prise de conscience de leur part, les sociétés de production semblent vouloir s'engager dans une évolution du cinéma. Malheureusement, tout désir d'évolution de la société s'heurte souvent à un rejet violent par une partie de la population. En me baladant sur les réseaux sociaux ces dernières années, il m'est alors apparu une augmentation des critiques négatives sur les films dits « masculins » mettant en avant essentiellement des femmes. Nous pouvons relever des exemples tels que *Ocean's 8* ou encore le reboot de *Ghostbusters* qui a notamment exaspéré Donald Trump. Ancrés dans leur fantasme de la femme, les hommes peinent à les imaginer autrement et crient rapidement au *pink-washing*. Malheureusement, les plateformes de *streaming* tendent à imposer leur vision, cherchant principalement le profit, quitte peut-être à négliger certains aspects de cet art qu'est le cinéma. Si le public est réticent envers des films plus égalitaires, cela pourrait alors à terme amener les plateformes à moins investir dans ces derniers. Viennent ainsi plusieurs questions : existe-t-il réellement une augmentation de la représentation féminine dans les films au cinéma ? Cette potentielle augmentation a-t-elle un effet sur les critiques du public ou les revenus engrangés par les films ? Pour répondre à ces questions, j'ai extrait près de 25000 films de la base de données du site internet The Movie Database (TMDb) à l'aide d'un API (**Fig. 1**). J'ai ainsi pu récupérer un maximum d'informations sur les différents films, telles que le casting, l'équipe technique, l'année, la note moyenne reçue par le film, etc.



**Figure 1. (A)** Liste des paramètres récupérés des sites TMDb & Bechdeltest.com via un API. **(B)** Liste de nouveaux paramètres, créés à partir des originaux, venant enrichir le dataset. Les noms en rouge correspondent aux variables qui n'ont pas été gardé par la suite.

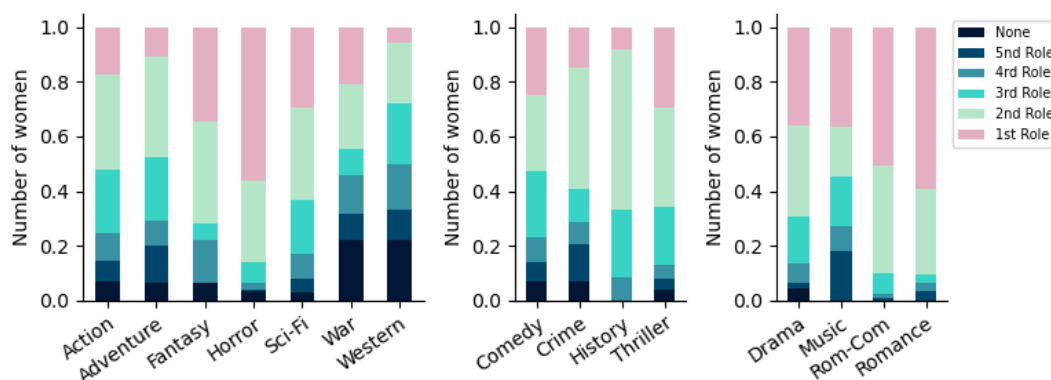
## ***Les femmes sont de plus en plus représentées... Mais restent minoritaires...***

Depuis les années 1970, nous observons une augmentation du nombre de premiers rôles féminins dans les films au cinéma qui a doublé en 40 ans (**Fig. 2**). Les films préférés par les femmes présentant la plus grande proportion de premiers rôles féminins. Pour autant les catégories dites « neutres » ou masculine présentent toujours près de 80% de films sans premier rôle féminin depuis 2010. Cette évolution des premiers rôles semble également s'accompagner d'une légère augmentation du nombre de femmes dans les films (**Fig. 2**). Cependant nous pouvons voir que le ratio reste situé en moyenne entre 1 et 2 femmes pour les 5 premiers personnages des films. Ce ratio dans les films « masculins » n'a pas sincèrement évolué avant les années 2000, mais évolue rapidement depuis.



**Figure 2.** Evolution du pourcentage de femme dans le premier rôle (à gauche) et du ratio de femme sur les cinq premiers personnages (à droite) en fonction des préférences de genre (Wühr, et, al, 2017).

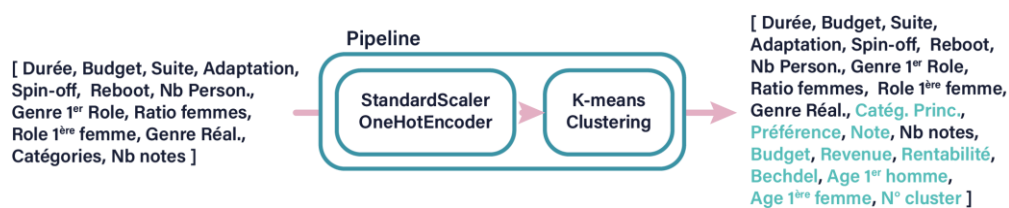
Depuis 2010, l'écrasante majorité des films présente au moins une femme. En moyenne, près de la moitié des films de chaque catégorie les mettent en avant dans le premier ou le second rôle (**Fig. 3**). Enfin, tandis que la majorité des catégories préférées par les hommes affichent un faible nombre de premiers rôles féminins, plus de 50% des films d'horreur sortis après 2010 présentent une femme dans le premier rôle (**Fig. 3**).



**Figure 3.** Importance du rôle du personnage féminin principal dans les différents genres de film. Les graphiques sont regroupés en fonction des préférences, les genres préférés par les hommes à gauche, les genres dits « neutres » au milieu et à droite les genres préférés par les femmes.

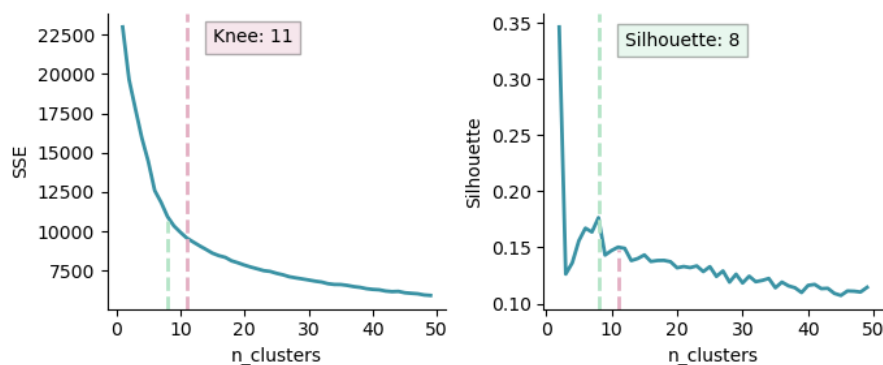
## Conséquence de la représentation féminine depuis 2010

Pour tenter de déterminer les profils des films les plus rentables j'ai décidé de regrouper les différents films sortis depuis 2010 à l'aide d'une méthode de *k-means* (**Fig. 4**). En effet, cette approche me permet d'associer les films en fonction de toutes les variables, non pas seulement basé sur la préférence de genre. Pour ce faire, j'utilise un dataset dépourvu des variables représentant une valeur d'appréciation du film (revenue, note, rentabilité). Ceci me permet de réaliser un *clustering* basé uniquement sur les variables les plus objectives (budget, nombre d'acteur, genre de film, etc). Ce dataset est alors envoyé dans un *pipeline* permettant la transformation des colonnes (*Scaler* et *OneHotEncoder*) puis la création des *clusters* (**Fig. 4**).



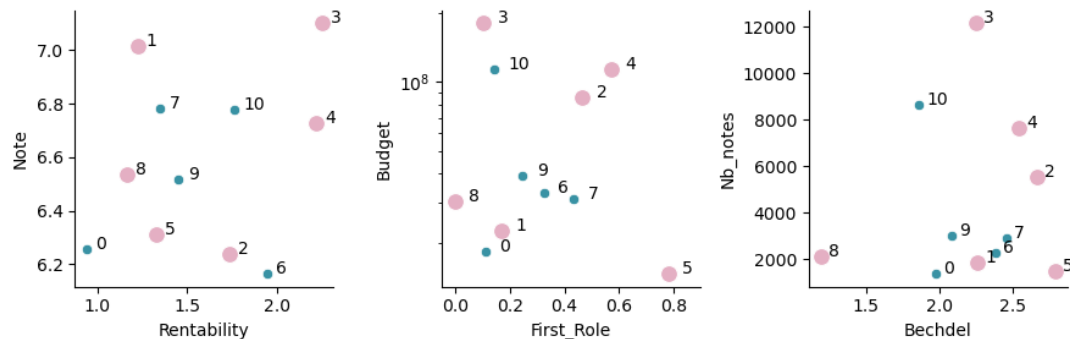
**Figure 4.** Liste des paramètres envoyés dans le pipeline permettant le clustering par la méthode des *k-means*. Un nouveau *DataFrame* est ensuite créé avec les paramètres ajoutés en bleu, dont le numéro de cluster dans lequel le film est placé.

Le nombre de clusters est déterminé à l'aide de la courbe de la somme des carrés des erreurs (SSE) et du score de silhouette. En comparant les deux courbes, je décide de récupérer le nombre de *clusters* donné par la localisation du coude de la courbe SSE grâce à la librairie python « kneed » (**Fig. 5**). J'ai ainsi pu déterminer que le nombre optimal est de onze *clusters*.



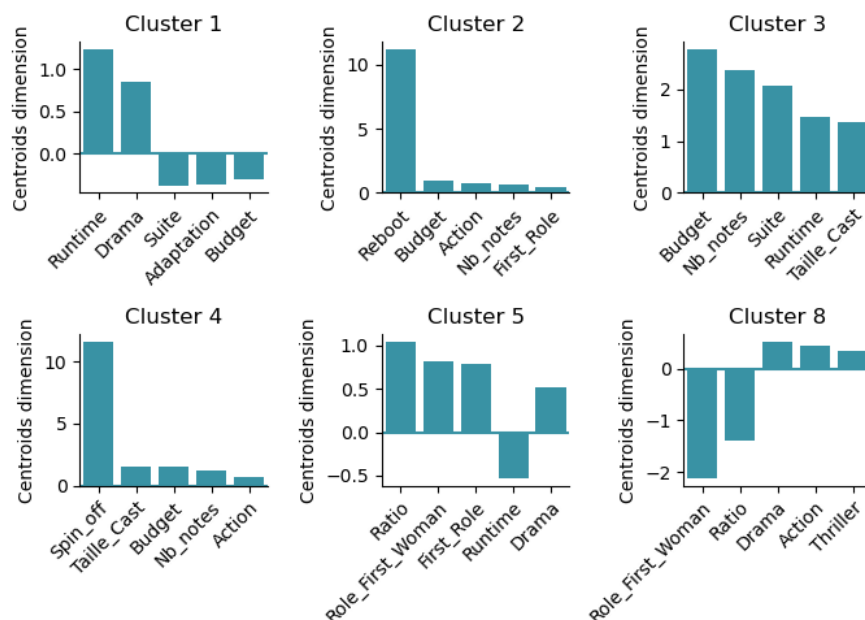
**Figure 5.** Représentation de la courbe SSE (à gauche) et du score de silhouette (à droite) en fonction du nombre de clusters. La valeur du genou de la courbe SSE est représentée en rose, la silhouette optimale est indiquée en vert.

Je porterai le reste de l'étude sur la moitié des *clusters* afin d'établir une vue d'ensemble sur l'importance de la représentation féminine dans la rentabilité, la popularité et le succès critique des films. Afin de les choisir, je projette le centre des clusters pour quelques variables (**Fig. 6**).



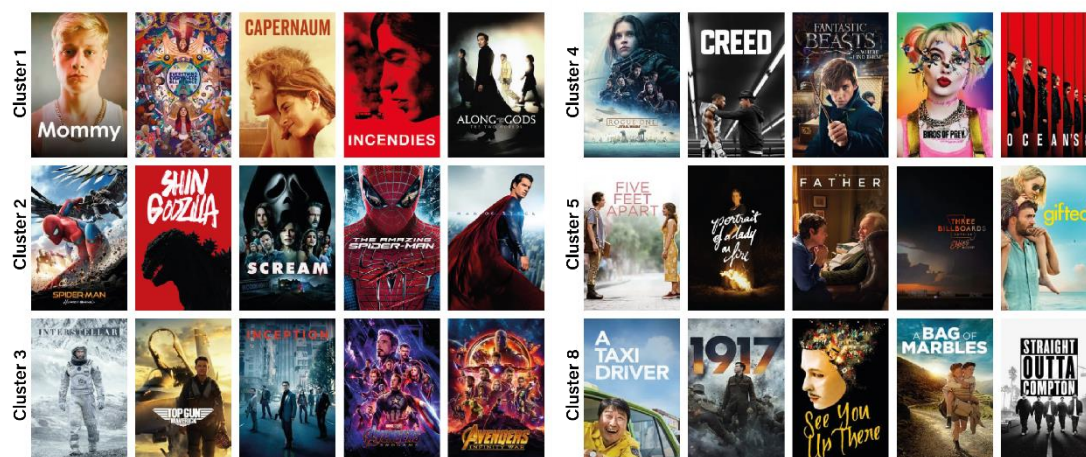
**Figure 6.** Projection des valeurs moyennes de chaque cluster pour différentes variables. Les clusters sélectionnés pour la suite de l'étude sont représentés en rose.

En analysant la position des centroïdes, j'ai pu ensuite extraire les variables les plus importantes pour la mise en place de chaque *cluster* (Y. Alghofaili, 2021). Cette analyse permet de donner leurs principales caractéristiques. Ainsi le *cluster 1* renferme plutôt des films dramatiques originaux à petit budget, le *cluster 2* des *reboot* de films d'actions et le *cluster 3* renferme des suites de film très populaires à gros budget avec beaucoup de personnages (**Fig. 7**).



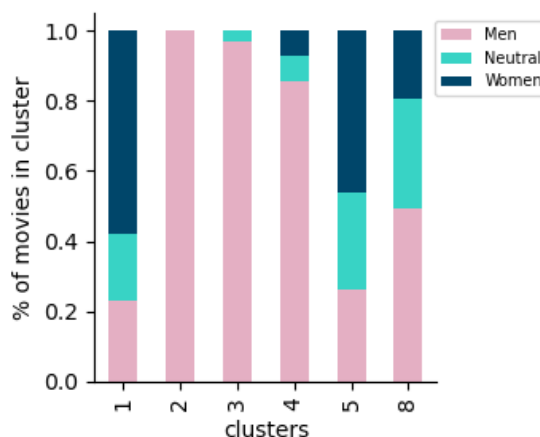
**Figure 7.** Représentation de l'importance des principales variables ayant permis de déterminer les différents clusters sélectionnés précédemment.

Je regarde ensuite quelques films présents dans ces *clusters* pour avoir une première vue des associations réalisées par le *clustering*. Cette observation basée sur les affiches est à mon sens positive et me satisfait quant au *clustering* réalisé (Fig. 8).



**Figure 8.** Exemples de films présents dans les six clusters sélectionnés.

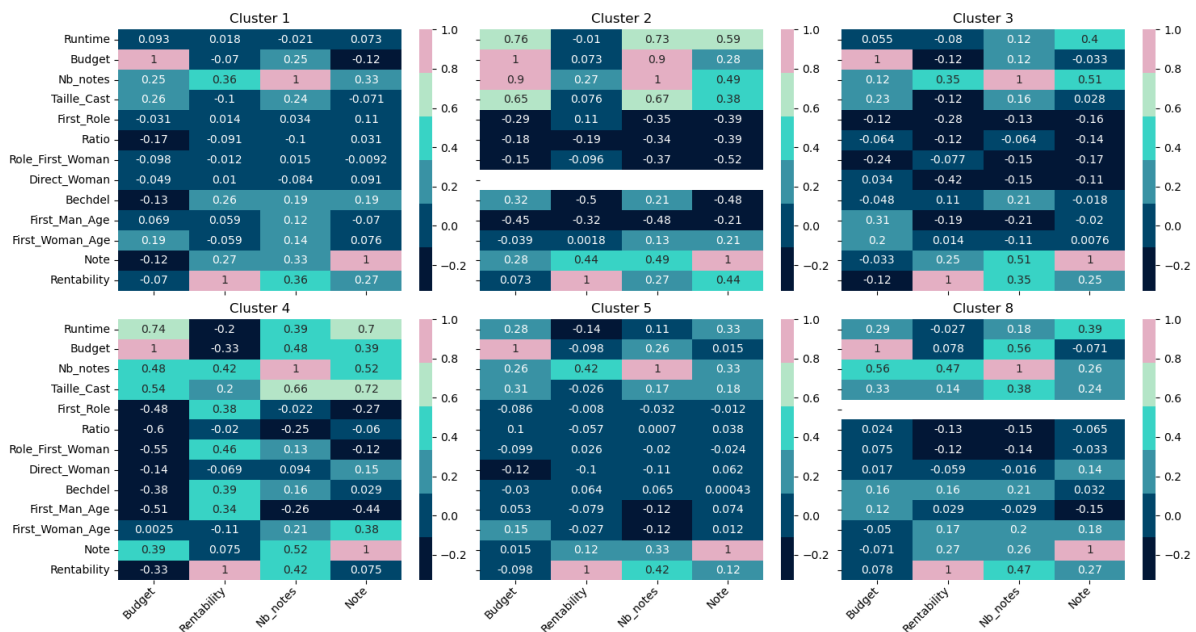
Nous pouvons notamment voir que le *cluster 3* qui est le plus rentable, le plus cher, le mieux noté et le plus populaire, est composé presque exclusivement de film dits « masculins » (Fig. 9) avec notamment les superproductions type « Avengers ». De l'autre côté nous avons par exemple le *cluster 1*, composé majoritairement de film dits « féminins » (Fig. 9) qui apparaît bien moins rentable, bien moins cher, bien moins populaire mais tout aussi apprécié par le public (Fig. 6). Pour finir je me suis penché sur chaque cluster pour déterminer



**Figure 9.** Proportion des catégories de films regroupées selon les préférences de genre dans les différents clusters.

d'éventuelles corrélations entre les différentes variables et les 4 paramètres qui me semblent les plus pertinents à savoir le budget, la rentabilité, la popularité donnée par le nombre de note et la note moyenne reçu par le film (Fig. 10). En effet, à mon sens il s'agit des principaux paramètres pouvant être pris en compte par les plateformes de *streaming* pour acheter, ou non, un film. Bien que cette approche ne me permette pas d'affirmer avec certitude mes observations, elle offre tout de même une piste de réflexion intéressante. La première chose qui saute aux yeux en étudiant les matrices de corrélations, c'est la forte influence des paramètres ayant attiré aux femmes dans les *clusters 2, 3, 4, et 8*, soit les *clusters* les plus « masculins » (Fig. 10).





**Figure 10.** Matrices de corrélation des différents clusters précédemment sélectionnés. En abscisse sont placées uniquement les variables d'intérêt pour la conclusion, à savoir le budget, la rentabilité, le nombre de notes et la note moyenne.

Le *cluster* 3, composé de films dits « masculins » à gros budget, semble être particulièrement affecté par la présence féminine. Ainsi il apparaît que les budgets alloués aux films les mettant en avant sont moindres, ces films semblent moins rentables, moins populaires mais également moins bien notés... De tels effets se retrouvent aussi un peu dans le *cluster* 8, où l'on peut voir que le nombre de femme à l'écran diminue la rentabilité et la popularité du film. Pour autant la réception critique, quant à elle n'y est pas influencée. Quant à eux, les *clusters* 2 et 4 décrivent un comportement plus contrasté. Tout comme dans le *cluster* 3, nous observons que la présence féminine induit des budgets plus faibles et malheureusement une mauvaise notation. Cependant, de manière surprenante nous constatons que dans le *cluster* 2, composé des *reboot*, la rentabilité ne subit pas sincèrement d'influence, malgré une popularité bien plus basse. Dans le *cluster* 4, composé des *spin-off*, la rentabilité est même plus haute pour les films mettant en avant une femme ! A l'opposé, les *clusters* 1 et 5, les plus féminins, ne sont pas influencés par le genre des personnages à l'écran.

## ***Le cinéma, représentation d'une société conservatrice ?***

Mon approche par *clustering* m'a permis d'aborder quelques points intéressants sur l'état du cinéma actuel. Ce regroupement de film par les *k-means* n'était peut-être pas la plus optimale pour ce problème, cependant l'utilisation de la méthode de *Spectral Clustering* ne m'a pas donné de résultats satisfaisants. Malgré cela, ce projet m'a permis de constater que la société, notamment les hommes, restaient très attachés aux archétypes de genre. Si les hommes préfèrent les films d'action c'est parce qu'ils y sont le plus représentés, de même que pour les femmes et les films dramatiques. Cependant là où les femmes sont habituées à voir des hommes jouer tous les rôles dans toutes les catégories de films, ces derniers semblent avoir des difficultés à voir des femmes jouer des rôles habituellement donnés aux acteurs. Ainsi la réception critique des films préférés par les hommes mettant en avant un personnage principal féminin a plus tendance à être négative. C'est le cas notamment des superproductions à gros budget dont l'analyse semble indiquer que ces films sont moins rentables que leurs homologues masculins. Si je reprends mon exemple du début, le film *Ghostbusters* de 2016, avec une note de 5.4, semble être considéré comme un navet. Pourtant il apparaît qu'il a engrangé 159% de bénéfice au box-office, alors pourquoi cette différence ? Simplement car contrairement aux entrées au cinéma, les critiques sur internet sont très majoritairement données par des hommes (89% selon le site IMDb). Des hommes qui se permettent également de noter le film pour surtout montrer leur rejet du projet, quand on peut voir que 35% d'entre eux ont mis la note minimale de 1/10 pour une note moyenne de 3.6. A l'inverse les femmes semblent se retrouver dans ce genre de film, puisque les notes données par les femmes sont bien meilleures avec une note moyenne pour *Ghostbusters* de 7.7/10 (**données issues d'IMDb**). Ce constat témoigne d'une nécessité d'augmenter le nombre de ce genre de projet féminins ! La popularité et la note moyenne reçu par un film sur internet ne devraient alors pas être un indicateur de réussite d'un film, car là aussi la vision masculine prime sur celle des femmes. Le futur du cinéma pourrait donc dépendre de l'importance donnée par les producteurs et plateformes de *streaming* à ces critiques. Il ne faudrait pour autant pas tomber dans l'extrême inverse et exagérer la présence féminine pour faire preuve de bonne foi face aux mouvements féministes. Dans ce sens, l'utilisation d'une notation telle que le *bechdel* pose problème à mon sens. Score de 0 à 3, ce test est aujourd'hui utilisé comme marqueur de sous-représentation des femmes, et donc indirectement comme marqueur féministe. Or il est aisé d'obtenir la meilleure note, sans améliorer pour autant la représentation de la femme au cinéma. Ce dernier semble être en crise, se contentant d'investir dans des *reboot*, des *spin-off* et multiples suites, qui sont aujourd'hui les plus rentables. Malheureusement le public, en particulier les hommes, ne souhaitent pas voir de femmes dans les rôles et les histoires auxquelles ils sont attachés.



L'acceptation des rôle féminins nécessitera encore un certain temps et surtout l'écriture d'aventures originales les sortants des archétypes habituels. Mais les réalisateurs en sont-ils capables ? Les plateformes leur en donneront-ils les moyens ou donneront-ils leur chance à plus de réalisatrices ?

## **Référence**

Tears or Fears? Comparing Gender Stereotypes about Movie Preferences to Actual Preferences (2017) – P. Wühr, B.P. Lange, S. Schwarz

Theatrical Market Statistics (2016) - Motion Picture Association of America

<https://towardsdatascience.com/interpretable-k-means-clusters-feature-importances-7e516eeb8d3c> (2021) - Y. Alghofaili

Lien vers le github : <https://github.com/Subnumine/DESU>