

# Analyse de données (Sociologie de la Culture)— L'influence entre les critiques

MEYNENT Fynch

Décembre 2023

## 1 Question de recherche

L'offre sur les marchés culturels étant strictement supérieure à la demande du fait du temps limité du public, la critique peut jouer un rôle essentiel dans l'établissement de la valeur des oeuvres culturelles. Pourtant, Beuscart J.-S. 2012 soulignent que l'effet de la critique n'est pas clair quand au succès populaire, bien que la critique et le succès populaire soit corrélés avec le budget dispensé dans l'oeuvre. L'attention médiatique aurait un impact davantage conséquent. On entend par la diffusion médiatique le fait que les individus et institutions sont informés de l'existence du film, la critique les jugements appliqués aux oeuvres, et le succès populaire le fait que les personnes aillent voir le film. Toutefois, Gemser G. 2006 montrent que dans le cadre du cinéma, la critique a un effet sur un certain public de taille plutôt restreinte : celui de l'art légitime.

Cependant, la critique en ligne apporte de nouvelles dimensions. En particulier, l'évaluation publique. Cela apporte de nouvelles dimensions à la critique : c'est ce que souligne Verboord 2014. Cela peut donc redonner une nouvelle dimension à la diffusion des hiérarchies culturelles ainsi qu'à la production de la valeur artistique. De plus, Faber J. 1984 souligne que les conseils interpersonnels ont un rôle important dans l'évaluation artistique que les individus portent à un film. Ainsi, cette évaluation publique pourrait potentiellement avoir cet effet de légitimité par les pairs. De plus, on pourrait aussi voir ces notes comme une autre indication du succès d'une oeuvre que le nombre d'entrée au Box Office.

Cependant, on sait que le succès populaire est aussi corrélé à l'impact médiatique. Ainsi, on ne peut pas non plus exclure le rôle d'acteurs d'influence plus large qu'un individu. Par suite, il serait intéressant d'investiguer le lien entre critique de la presse et critique des spectateurs pour comprendre comment s'intriquent ces deux tendances.

Dans quelle mesure la critique médiatique impacte-t-elle l'évaluation en ligne effectuée par le public ?

## 2 Données et méthode d'analyse

### 2.1 Description de la base de données

Pour produire les analyses que l'on va détailler d'un second temps, on va s'appuyer sur une base de données issue d'Allociné. Elle contient entre autre, pour 102910 films, l'année de production, une note moyenne attribuée par la presse, une note moyenne attribuée par les spectateurs, le budget, la nationalité et la présence de récompense. En retirant les films ne présentant pas de note de presse ou de notes de spectateurs, on a alors plus que 12647 films. Cela tendra à légèrement surreprésenter certains films : ceux étant plurilingues. (le traitement précis des valeurs manquantes est évoqué en annexe).

Les notes sur Allociné prennent des valeurs entre 0.5 et 5. La figure 1 nous donne une idée de la répartition des valeurs.

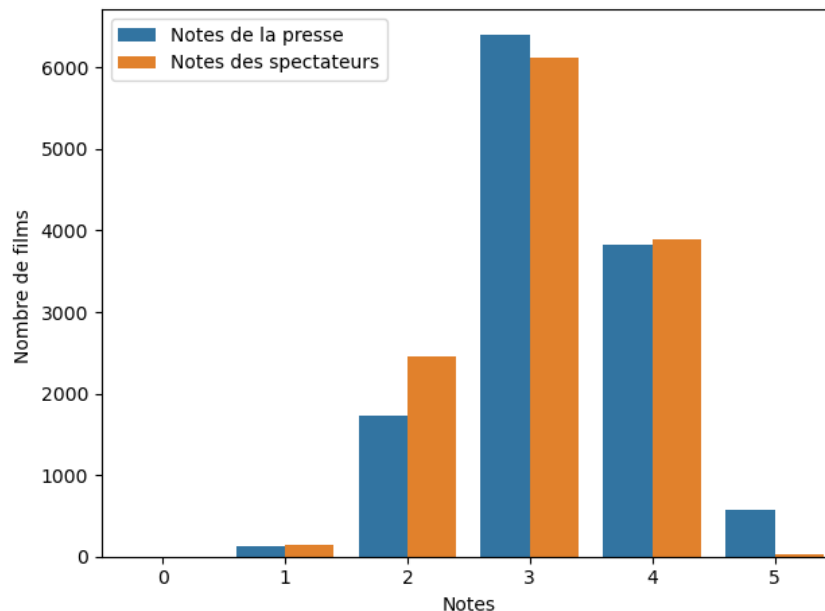


Figure 1: Répartition des notes

Cela nous permet d'observer que les notes semblent relativement corrélées. En effet, la répartition est proche. Cependant, les spectateurs semblent noter plus durement que la presse.

## 2.2 Méthode d'analyse

Afin d'établir les effets des notes de la presse sur les notes des spectateurs, nous allons effectuer la régression linéaire suivante :

$$note_{spectateurs} = \alpha + \beta note_{presse} + \epsilon \quad (1)$$

Cependant, on peut supposer que le fait qu'un film soit récompensé a aussi un impact sur les notes. Alors, nous allons tester la corrélation pour voir si cela est un facteur confondant. Aussi, rappelons que Beuscart J.-S. 2012 montre que le budget est un facteur confondant dans la corrélation entre succès populaire et critique. On va donc effectuer des régressions en contrôlant par ces deux éléments. Enfin, un dernier facteur que nous allons utiliser comme variable de contrôle est le nombre d'entrées au box-office en France qui servira de mesure de succès populaire.

## 3 Résultats

La corrélation entre note de presse et note des spectateurs est indéniable. En effet, on voit dans la table 1 qu'on peut rejeter l'hypothèse nulle au seuil de 1%.

	Coefficient	Std Err	t	p-value	[0.025	0.975]
<b>Constante</b>	1.2913	0.025	50.924	0.000	1.242	1.341
<b>Note Presse</b>	0.5478	0.008	70.657	0.000	0.533	0.563

Table 1: Résultat de la régression sans contrôles

On a  $R^2 \simeq 0.304$

En revanche, on s'aperçoit que la note n'est pas corrélée à la récompense (Table 3 en annexe). Pourtant, on sait que certaines institutions ont un certain prestige. Par exemple, Thévenin 2021 dit que le festival de Cannes joue un rôle non négligeable dans la valorisation des oeuvres cinématographiques ainsi que dans leur diffusion sur les marchés artistiques. Ainsi, on peut supposer qu'il y'a probablement une hétérogénéité des récompenses et distinction. Cependant, on ne peut pas le mesurer ici du fait des limites dans la base de données qui nous indique seulement si oui ou non un film a reçu une quelconque distinction. Egalement, le budget n'est pas corrélé à la note de la presse (Table 4 en annexe). Cela peut sans doute s'expliquer par le fait que nous avons exclu dans les valeurs manquantes des films probablement moins diffusés, donc avec moins de budget, qui n'ont pas pu recevoir de notes de la presse. Ainsi, il y'a un probable biais de sélection.

En revanche, le nombre d'entrées au box office est statistiquement significatif dans la régression avec contrôle. En effet, nous voyons dans la table 2. Les variables ont été standardisées afin de pouvoir comparer les ordres de grandeur.

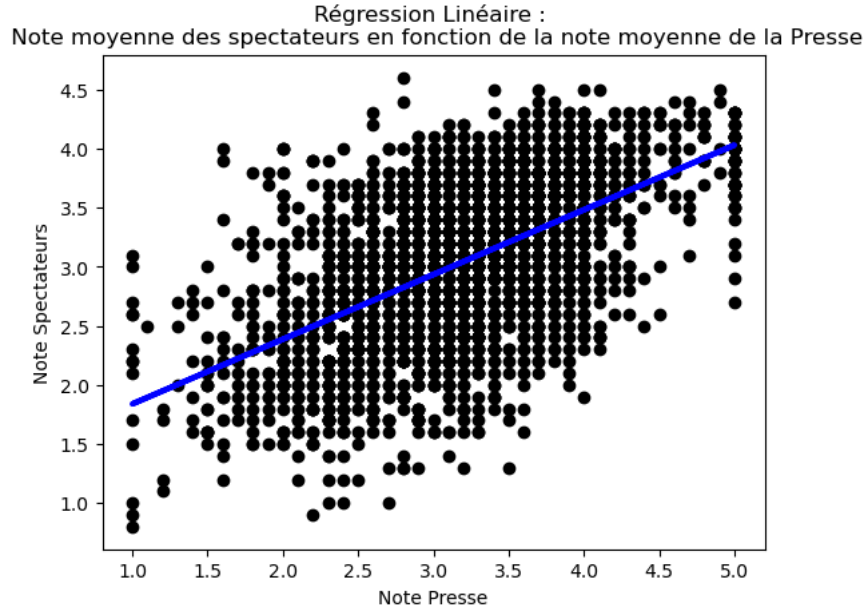


Figure 2: Régression de la note des spectateurs sur la note de la presse

	Coefficient	Std Err	t	P t	[0.025	0.975]
<b>Constante</b>	3.0047	0.007	432.284	0.000	2.991	3.018
<b>Note Presse</b>	0.3725	0.007	53.355	0.000	0.359	0.386
<b>Nombre d'entrées</b>	0.0669	0.007	9.576	0.000	0.053	0.081

Table 2: Résultats de la régression contrôlée par le nombre d'entrées au Box Office français

Ainsi, il apparaît que toute chose égale par ailleurs, les notes de la presse semblent tout de même davantage corrélées que le nombre d'entrées au Box Office. Il est difficile d'identifier clairement un effet causal car on ne peut pas être certain que des facteurs confondants ne biaisent pas nos résultats.

## 4 Conclusion et limite

Les notes des spectateurs sont probablement influencées par la presse. Cependant, les mécanismes à l'oeuvre sont inconnus, d'autant plus que certains résultats sont étonnants en terme de récompense et de budget.

Ainsi, ce travail présente de nombreuses limites. Tout d'abord, il suppose une homogénéité des publics et de la presse. Or, des registres de légitimité culturelle et des logiques de distinction (Bourdieu 1979) à l'oeuvre montrent que ce n'est pas le cas. Ensuite, il aurait été intéressant d'intégrer d'autres variables

de contrôle que nous n'avons pas à disposition. Par exemple, Glenski M. 2017 identifie des effets de mimétisme : une manipulation aléatoire des réactions sur les réseaux sociaux montre que des effets de mimétisme positif ou négatifs sont à l'oeuvre dans les réactions. Dès lors, les premières notes données sur AlloCiné risquent d'influencer les suivantes. Ainsi, il aurait été intéressant de pouvoir contrôler par la valeur des premières notes de spectateurs. On aurait pu également chercher à contrôler par : le nombre de projections ou de cinémas où sont diffusés les films, le nombre de semaines en salle, la diffusion par DVD ou streaming etc. afin d'identifier les biais de note liés aux différences de diffusion. Enfin, notre échantillon, de part ces valeurs manquantes, tend à surreprésenter des films bénéficiant d'une diffusion plus large.

## References

- Beuscart J.-S., Mellet K. (2012). “Promouvoir les œuvres culturelles. Usages et efficacité de la publicité dans les filières culturelles”. In: *Ministère de la Culture*.
- Bourdieu, P. (1979). *La Distinction, Critique sociale du jugement*. Editions de Minuit.
- Faber J., O’Guinn T. (1984). “Effect of Media Advertising and other Sources on Movie Selection”. In: *Sage*.
- Gemser G. Oostrum M. V., Leenders M. A. A. M. (2006). ““The impact of film reviews on the box office performance of art house versus mainstream motion pictures””. In: *Journal of Cultural Economics* 1.31.
- Glenski M., Weninger T. (2017). “Rating Effects on Social News Posts and Comments”. In: *ACM Transactions on Intelligent Systems and Technology* 8.
- O’Sullivan, C. (2011). *Translating Popular Film*. Palgrave MacMillan.
- Thévenin, O. (2021). “le Festival de Cannes et la montée du cinéma d’auteur à l’ère de la mondialisation”. In: *Loisir et Société*.
- Verboord, M. (2014). “The impact of peer-produced criticism on cultural evaluation: A multilevel analysis of discourse employment in online and offline film reviews”. In: *New Media Society* 6.16.

	Coefficient	Std Err	t	p-value	[0.025	0.975]
<b>Note Presse</b>						
Constante	3.1935	0.008	413.343	0.000	3.178	3.209
Récompense	-0.0170	0.020	-0.847	0.397	-0.056	0.022
<b>Note Spectateurs</b>						
Constante	3.1935	0.008	413.343	0.000	3.178	3.209
Récompense	-0.0170	0.020	-0.847	0.397	-0.056	0.022

Table 3: Résultats de la régression des notes sur la Récompense

## A Résultats des régressions avec variables de contrôle

On a implémenté les deux modèles suivants :

$$note = \alpha + \beta_1 récompense + \epsilon \quad (2)$$

$$note = \alpha + \beta_1 budget + \epsilon \quad (3)$$

On obtient les résultats suivants :

Notes Spectateurs	Coefficient	Std err	t	p-value
Constante	3.1403	0.017	181.014	0.000
Budget	$1.392 \times 10^{-12}$	$1.17 \times 10^{-11}$	0.119	0.905
<b>Notes Presse</b>				
Constante	3.1254	0.019	165.967	0.000
Budget	$1.194 \times 10^{-11}$	$1.27 \times 10^{-11}$	0.940	0.347

Table 4: Résultats de la régression des notes sur le Budget

On voit bien la corrélation de ces variables n'est pas certaine car on ne peut rejeter l'hypothèse nulle au seuil de 10%.

## B Davantage de statistiques descriptives

### B.1 Analyse des valeurs manquantes

Pour vérifier la répartition des valeurs manquantes, on va contrôler par une variable non-utilisée dans nos modèles qui est le nombre de langues utilisées dans le film. Malheureusement, je n'ai pas réussi à trouver, dans la littérature, des articles parlant de l'impact du multilinguisme dans la diffusion des oeuvres en terme quantitatif. Cela restait plutôt dans des thématiques de la représentation du plurilinguisme au cinéma (par exemple, dans O'Sullivan 2011). J'ai donc supposé que le plurilinguisme pouvait être un critère impactant légèrement la diffusion du film et pouvant être révélateur d'un certain type d'histoires où

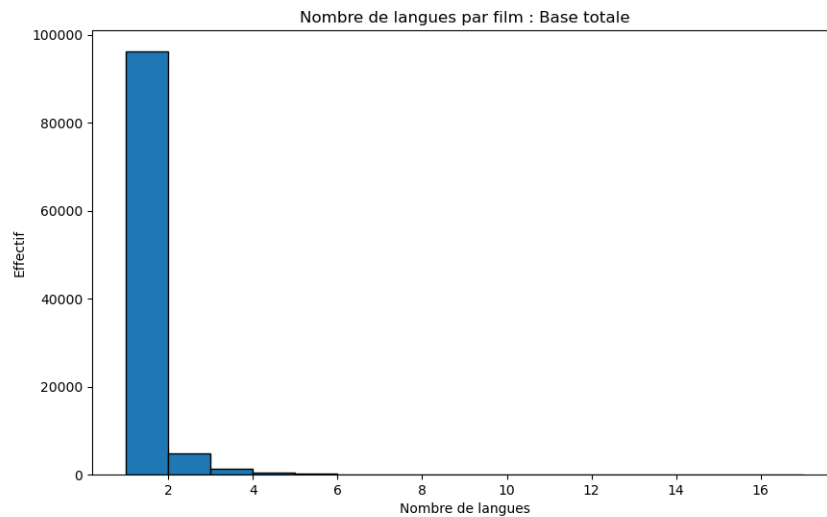


Figure 3: Répartition du nombre de langues disponible pour le film

plusieurs langues sont nécessairement représentées. C’est par exemple le cas dans un film traitant d’immigration, mais ça l’aït moins dans certains blockbuster de fiction.

### B.1.1 Analyse des valeurs manquantes pour les notes

Notre sous-échantillon a tendance a surreprésenter les films plurilingues.

### B.1.2 Analyse des valeurs manquantes pour le budget et le box Office

Notre sous-échantillon avec les variables de contrôle a tendance a légèrement surreprésenter les films plurilingues par rapport à notre sous-échantillon initial.

## B.2 Plus d’informations sur le corpus

Voici quelques informations supplémentaires pour la répartition des valeurs de nos variables d’intérêt.



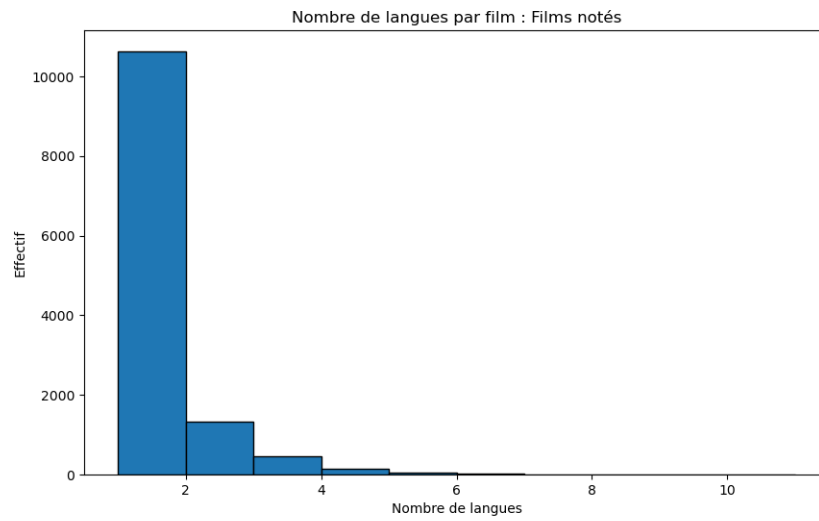


Figure 4: Répartition du nombre de langues disponible pour le film sur l'ensemble des films comprenant des notes

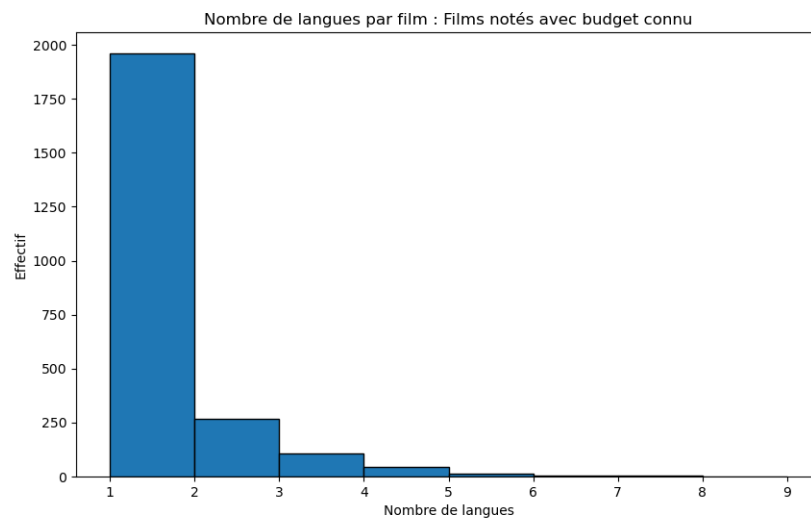


Figure 5: Répartition du nombre de langues disponible pour le film sur l'ensemble des films comprenant des notes et un budget explicité

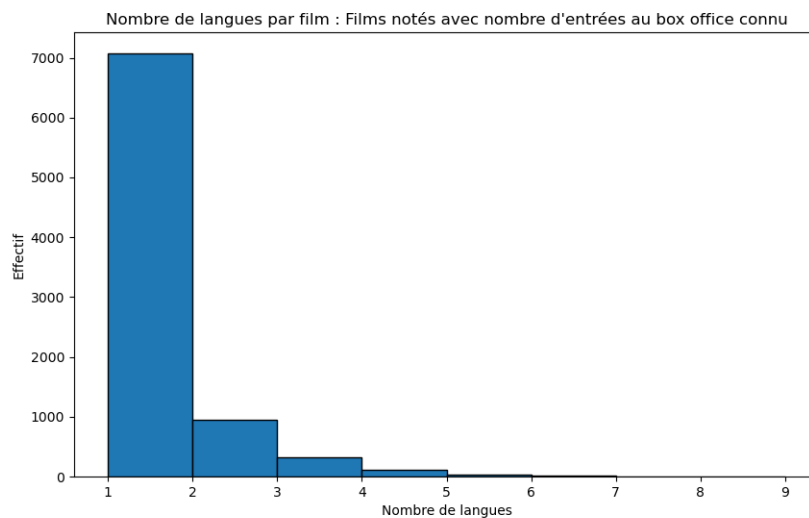


Figure 6: Répartition du nombre de langues disponible pour le film sur l'ensemble des films comprenant des notes et un nombre d'entrée au box-office explicité

	Note Presse	Note Spectateurs	Année de production	Récompense
mean	3.189215	3.040650	2006.453136	0.148652
std	0.718512	0.684756	14.775440	0.355759
min	1.000000	0.800000	1921.000000	0.000000
25%	2.700000	2.600000	2002.000000	0.000000
50%	3.200000	3.100000	2009.000000	0.000000
75%	3.600000	3.600000	2016.000000	0.000000
max	5.000000	4.600000	2023.000000	1.000000

Table 5: Statistiques descriptives des principales variables d'intérêt