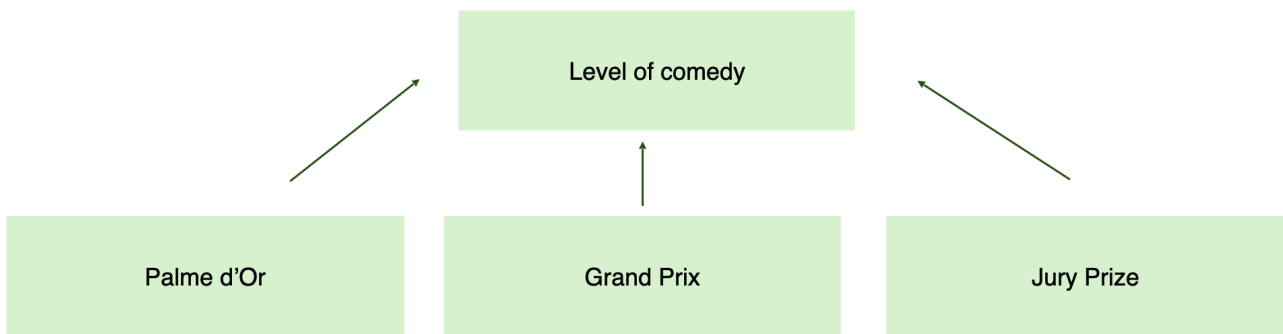


Is There a Disdain of Comedy in the Festival de Cannes?

Author : Noam Chabot
Date : January 2025

Is there a bias against comedy in the selection process of the Cannes Festival?

To answer this question, we formulated three hypothesis:



Predictions:

- **P1** - The winners of the Palme d'Or tend to have a lower level of comedy than the official Selection.
- **P2** - The winners of the Grand Prix tend to have a lower level of comedy than the nominated films.
- **P3** - The winners of the Jury Prize tend to have a lower level of comedy than the nominated films.

Data

We constructed the dataset using the Wikipedia API.

We first used the Wikipedia MediaWiki API to check, for each year since the creation of the Festival de Cannes (i.e., from 1946 to 2025), whether a page with the exact title “YYYY Cannes Film Festival” exists, automatically following redirects.

When such a page is found, the main metadata (year, canonical title, page ID, and full URL) are collected into a structured list. The only two years for which no Wikipedia page exists are 1948 and 1950, which is consistent with the fact that the festival was cancelled in those years.

We then scraped, for each existing page, the full list of films competing in the Official Competition. The Official Competition should be distinguished from the Official Selection, which includes all films chosen by the festival to be presented under its official label and encompasses several distinct sections, such as the Official Competition, Un Certain Regard, and Out of Competition screenings. While all these films are part of the Official Selection, only those in the Official Competition are eligible for the Palme d’Or. The Official Competition is therefore a subset of the Official Selection, restricted to films competing for the main awards, including the Palme d’Or. We chose to focus on this category because it is the most prestigious and the one most prominently highlighted in film promotion and media coverage.

We then enriched each film with a synopsis through additional scraping, as well as with metadata such as the director, the genre as provided by Wikipedia, and the film’s country of origin using the API. After data processing and cleaning, the resulting dataset contains a total of $N=1,799$ films.

We then created three additional datasets by scraping the respective Wikipedia pages that list all winners in the Palme d’Or, Grand Prix, and Jury Prize categories. These three categories were selected because they are considered major awards and are attributed to films rather than to individuals, unlike, for example, the Best Director Award. The Palme d’Or represents the festival’s highest distinction and is awarded to the best film in the Official Competition. The Grand Prix is the second most important award, while the Jury Prize is a collective distinction recognizing a film for its originality or artistic singularity. We did not include the Best Director Award, as our analysis focuses on films themselves and on the overall cinematic work presented to audiences, rather than on the individual director. All the manipulation and scraping on Wikipedia had been done on 6th January 2026.

Finally, we concatenated these three sub-datasets with the main dataset using the film’s Wikipedia URL as the key, as this link is unique. We removed any rows corresponding to the years 1948 and 1950. We then addressed an important specificity of the Festival de Cannes: the major awards have not been awarded every year since the festival’s inception. For instance, the Grand Prix in its current form—that is, as the second most important award after the Palme d’Or—was introduced in 1967, and the Jury Prize has not been awarded in some years. Consequently, we created three variables, Palme d’Or, Grand Prix, and Jury Prize, which take the value *Official Competition* if the film competed in that category, the name of the prize if the film won it, and *NaN* if the prize was not awarded in that year.

We therefore obtain a final dataset that includes all films selected for the Official Competition and indicates, for each major category, whether the film won the award or not, with the constraint that a film cannot win more than one major prize, as the competition rules impose exclusivity among major awards. The dataset is going to be published on Kaggle, as no existing dataset provided a comprehensive and well-structured record of all films selected for the Official Competition, and even the few datasets listing Palme d’Or winners were poorly formatted.

For the purposes of our specific project, we decided to exclude films selected in 1968 and 2020, as no awards were granted in those years due respectively to the events of May 1968 and the COVID-19 pandemic. We also removed films for which genre information was missing, as this variable is central to the comparisons conducted in our analysis.

Finally, we created a new variable measuring the “level of comedy” of each film, based on the Wikipedia genre tags, using the following algorithm:

- **0 (No Comedy)** – The Comedy tag is absent from the list of genres.
- **1 (Mixed Comedy)** – The Comedy tag is present alongside at least one of the following genres: Drama, Action, Romance, or Horror.
- **2 (Pure Comedy)** – The Comedy tag is present, and none of the four genres above appear.

These four genres were chosen because, in classical and neoclassical theatre traditions, they were considered incompatible with Comedy. For example, a comedy could not include elements of horror. The case of Romance is more nuanced: romantic comedies existed in classical theatre, but from a contemporary perspective, romantic comedy and pure comedy slightly differ in both narrative and humorous structure.

The final dataset included **N = 1,802 films**.

We used three main variables of different types to test our predictions:

- Ordinal: the **level of Comedy** (see the scale above, **Figure 2a**);
- Binary: the film’s status in the **Best Picture** category (Winner or Nominated, **Figure 2b**)
- Binary: the film’s status in the **Best Director** category (Winner or Nominated, **Figure 2c**)
- Binary: the film’s status in the **Best Director** category (Winner or Nominated, **Figure 2c**)

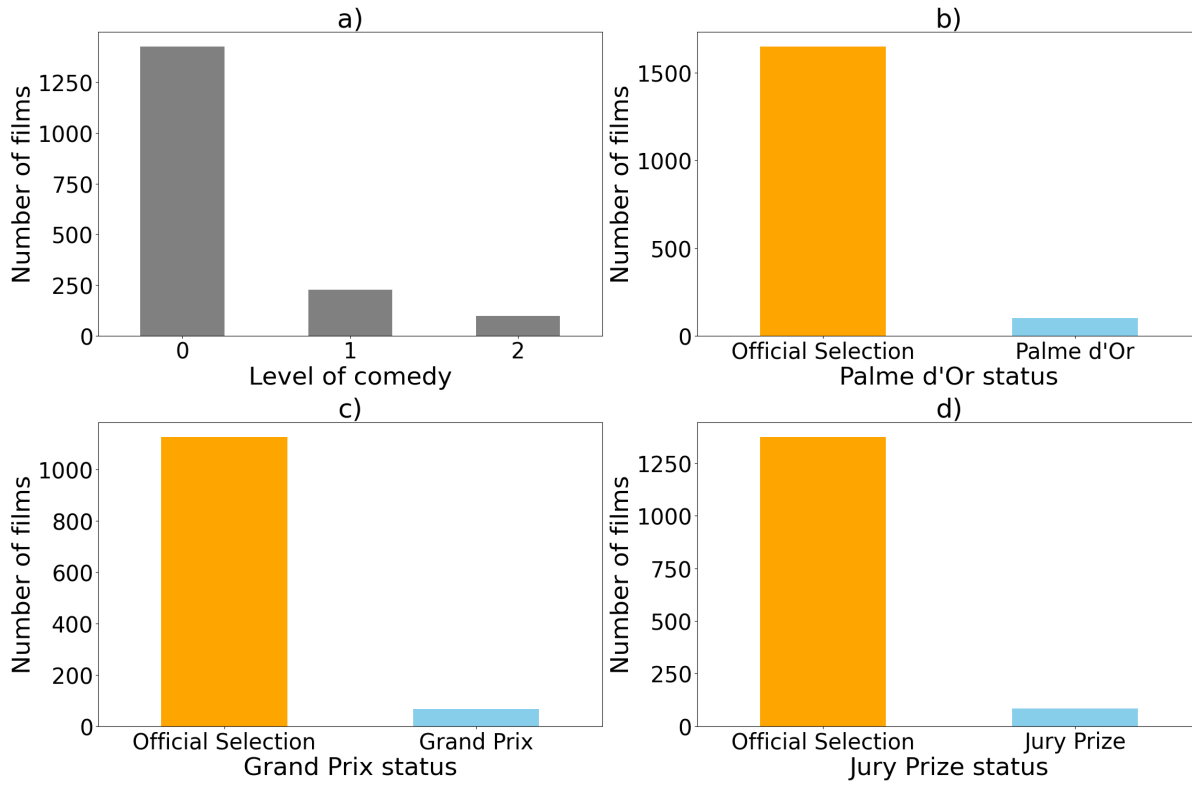


Figure 2 - Repartition of films according to a) the level of comedy (with 0: No Comedy, 1: Comedy mixed with another genre, 2 : Pure comedy) and the film status at the Festival de Cannes for b) the Palme d'Or, c) the Grand Prix and d) the Jury Prize

Methods

To test our three predictions, like the César Award, we compared the level of comedy in movies depending on whether they won the Palme d'Or, Grand Prix or Jury Prize, or not (**Figure 3.d, 3.e and 3.f**). We conducted two independent **t-tests**, assuming equal variance and an equivalent gap between categories 0 and 1, and between 1 and 2. The corresponding formula is:

$$t = \frac{\mu_1 - \mu_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Where :

μ_1 and μ_2 : mean comedy levels of each group

s_p^2 : pooled variance

n_1 and n_2 : sample sizes of the two groups

Results

The mean of the comedy level of films among the winners of the best Picture is slightly higher than for the nominated ($\mu_1 = 0.218$, $\mu_2 = 0.242$, see **Figure 3.e**), but this difference is not significant ($t = 0.523$, $p = 0.602$). Plus, the mean of the comedy level of movies among the winners of the Grand Prix is also slightly lower than for the nominated ($\mu_1 = 0.227$, $\mu_2 = 0.242$, see **Figure 3.f**), but this difference is not significant neither ($t = 0.233$, $p = 0.816$). To finish, the mean of the comedy level of films among the winners of the Grand Prix is slightly higher than for the nominated ($\mu_1 = 0.259$, $\mu_2 = 0.249$, see **Figure 3.g**), but this difference is not significant neither ($t = -0.175$, $p = 0.862$)

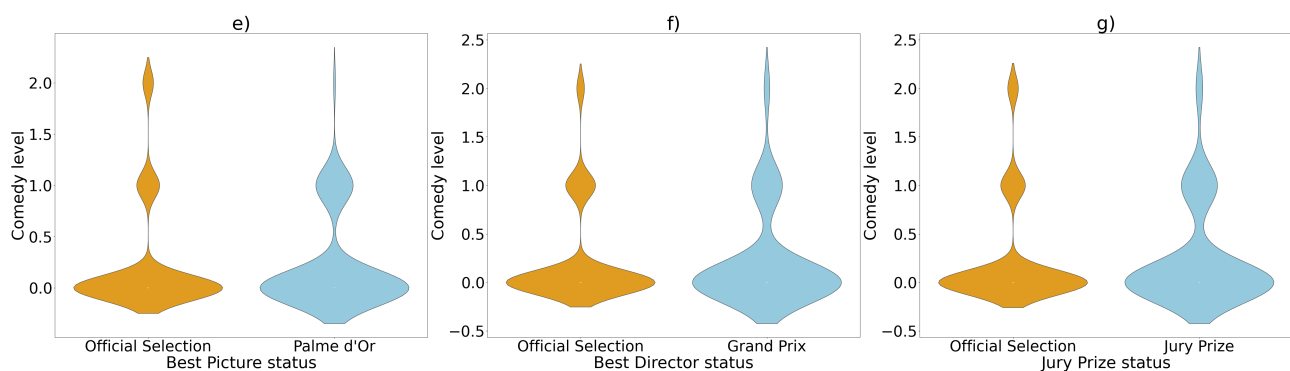


Figure 3 - Repartition of the level of comedy (with 0: No Comedy, 1: Comedy mixed with another genre, 2 : Pure comedy) according to e) the Palme d'Or status, c) the Grand Prix status and d) the Jury Prize status

Conclusion

This study set out to test whether comedy is systematically disadvantaged at the Cannes Film Festival at the stage of award attribution. Across the three major prizes considered—the Palme d'Or, the Grand Prix, and the Jury Prize—we found no statistically significant differences in the level of comedy between winning films and nominated films. Although minor variations in mean comedy levels were observed, these differences were small, inconsistent across awards, and well within the range of statistical noise. Taken together, the results do not support the hypothesis of a measurable anti-comedy bias in the final selection of winners at Cannes.

However, the absence of significance at the award stage does not rule out the existence of bias elsewhere in the festival's decision-making process. An important avenue for future research lies

upstream, in the composition of the Official Competition itself. If comedies are underrepresented, or systematically classified as mixed genres before reaching the nomination stage, the potential bias may operate more strongly in the selection of nominated films than in the choice of winners among them. In this sense, Cannes may not so much “disdain” comedy when awarding prizes as filter it out earlier, shaping the pool of contenders in ways that remain largely invisible when focusing on winners alone.