

Boostrapping with French Cinema Awards



Predictions:

- **P1** - Films nominated for the Cannes Film Festival exhibit a lower level of comedy than the full set of eligible films
- **P2** - Films nominated for the César Awards exhibit a lower level of comedy than the full set of eligible films

Data

With regard to film festivals, we reused the same data previously collected for the Cannes Film Festival and the César Awards. However, we reintegrated films nominated in 2020 and 1968 for the Cannes Film Festival, since our analysis focuses exclusively on nominees rather than winners; these years are therefore relevant to our hypothesis.

As a comparison baseline, we chose to use the TMDB dataset, which contains $N = 1,306,063$ films in its raw, unmanipulated form. The main challenge lies in identifying all films eligible for both festivals, as their eligibility criteria are not exactly the same.

First, we processed the TMDB dataset to match the criteria common to both festivals. We removed films labeled with the genre “Documentary” as well as those marked “yes” in the *Adult* column, as these types of films are not typically nominated at festivals. We also restricted the dataset to films with a runtime exceeding 60 minutes, an explicit requirement of both festivals for a film to be considered a feature-length production.

Moreover, films selected at Cannes or nominated for the César Awards typically have a minimum level of industrial existence and receive a theatrical release (at least in France after selection). As a result, they are generally covered by the specialized press. We therefore assume that a film selected at Cannes has received at least 10 votes on TMDB. Films with fewer than 10 votes are often titles that were never truly released, or that correspond to amateur or local productions, television films, or database errors. In all cases, such entries do not match the profile of films selected at Cannes.

We therefore restrict the comparison set to films that have received at least 10 votes on TMDB, in order to exclude entries lacking minimal distribution or public visibility. This criterion allows us to better approximate the pool of films eligible for Cannes and the César Awards, without introducing a strong bias toward widely successful films. Finally, we further restricted the dataset to films that were actually released theatrically.

Cannes-specific criteria.

Feature-length films submitted for selection at the Cannes Film Festival must in particular meet the following requirements: they must have been produced within the twelve months preceding the festival; they must not have been commercially exploited outside their country of origin; they must not have been presented at another international film festival; and any feature-length film invited to compete must receive a commercial theatrical release in France, in compliance with French regulations, particularly those governing release windows. Unfortunately, it is difficult to verify that all of these criteria are fully satisfied using our available data.

As a pragmatic solution, we focused on the production country. We retained only the first listed country in cases of co-production. This choice mirrors our Cannes dataset, which also sometimes includes multiple countries, and simplifies alignment across datasets. We assume that co-producing countries are listed in the same order in both datasets, with the first country corresponding to the principal producer. After applying these restrictions, the resulting dataset contains $N = 59,913$ films (see **Figure 1.a, 1.b, 2.e and 2.f**).

César-specific criteria.

A feature-length film is eligible for the César Award for Best Film if it meets all of the following conditions: it must be a feature-length film of French initiative (i.e., majority or relative-majority French production, regardless of shooting language); it must have received CNC production approval; it must hold a theatrical exhibition visa issued by the French Ministry of Culture (excluding temporary or exceptional visas); it must have had its first commercial theatrical release in at least one commercial cinema in the Paris region between January 1 and December 31 of the year preceding the ceremony; and it must have been commercially exhibited for at least one week (seven consecutive days) in that cinema.

Once again, we do not have access to all of this information within the TMDB dataset. However, unlike Cannes, where films may be international, we are able to account for French production in the case of the César Awards. We therefore restricted the dataset to films whose country of production is France. Because films must be primarily produced in France, we again retained the first listed country in cases of co-production, treating it as the principal producer. After applying these constraints, the final dataset contains $N = 5,187$ films (see **Figure 1.a, and 1.b**).

These operational choices reflect pragmatic approximations necessitated by data availability constraints in TMDB. While several official eligibility criteria for Cannes and the César Awards cannot be directly verified, the applied filters are intended to approximate as closely as possible the

institutional selection pools of both festivals. Importantly, these restrictions are symmetrical with respect to genre and do not specifically disadvantage comedy, thereby limiting the risk of introducing systematic bias. As such, the resulting datasets should be interpreted as conservative but plausible representations of the films realistically eligible for selection at Cannes and the César Awards.

To measure the level of comedy, we chose to use the mean value of the comedy metric ranging from 0 to 2 that was already implemented (see **Figure 3.g, 3.h, 3.i, 3.j**). We therefore assume that a pure comedy is “twice as comedic” as a mixed-genre comedy. This assumption was adopted pragmatically in order to simplify the calculations and to rely on a single statistical test, and it is justified in the same manner as in our previous analyses of the Cannes Film Festival and the César Awards. In addition, this approach is considerably easier to implement than an LLM-based method when dealing with a large number of films, as is the case here.

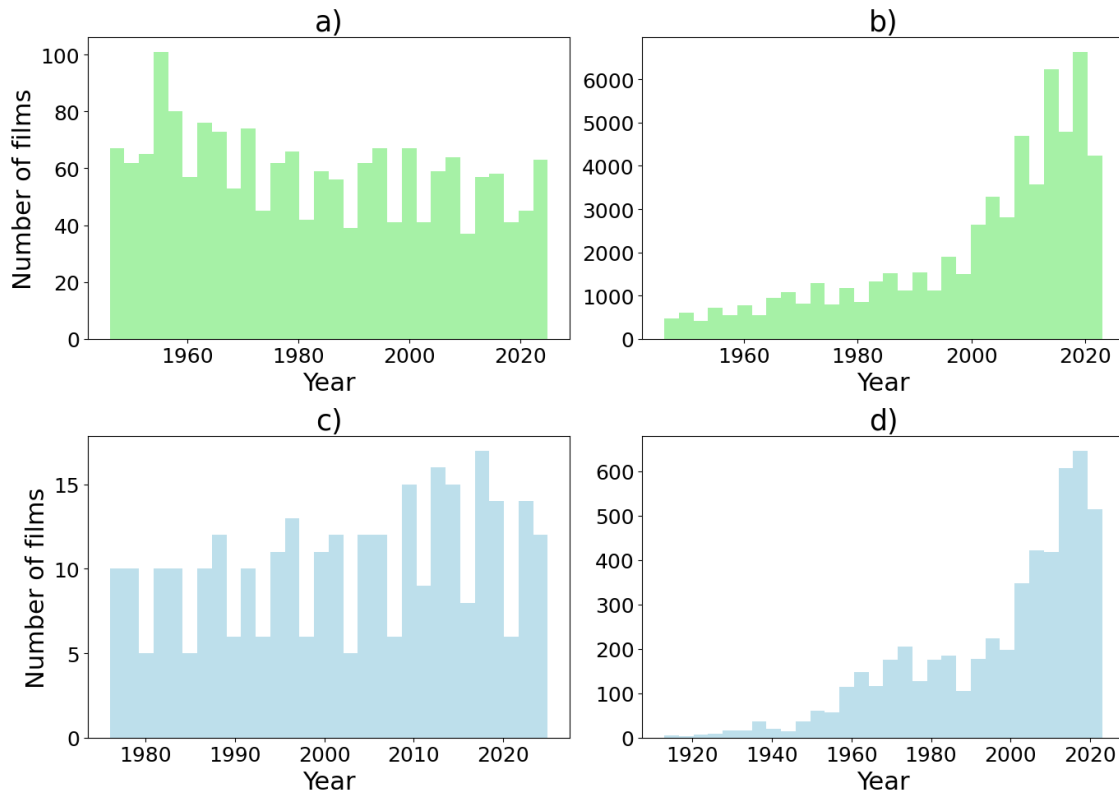


Figure 1 - Distribution of films by release year for a) César Award nominees, b) the corresponding TMDB comparison sample, c) Cannes Film Festival selections, and d) the corresponding TMDB comparison sample.

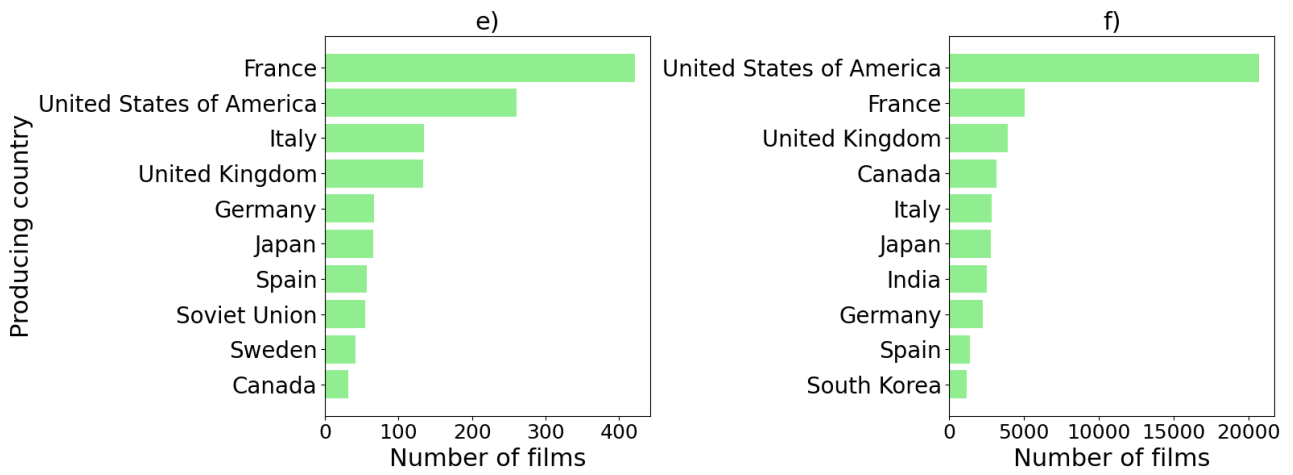


Figure 2 - Distribution of the ten most frequent producing countries for e) César Award nominees and f) the corresponding TMDb comparison sample

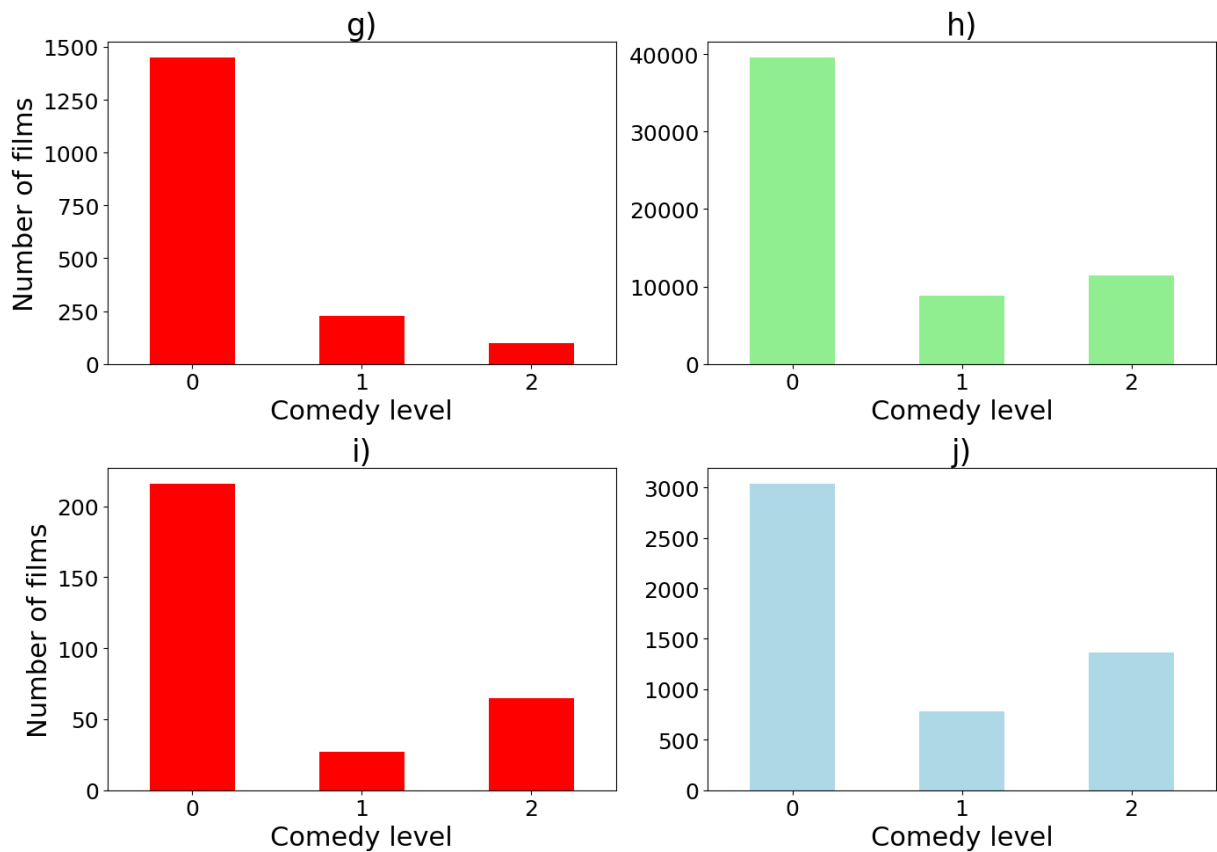


Figure 3 - Distribution of films by comedy level (0 = no comedy, 1 = mixed-genre comedy, 2 = pure comedy) for g) César Award nominees, h) the corresponding TMDB comparison sample, i) Cannes Film Festival selections, and j) the corresponding TMDB comparison sample.

Method

To test whether the official selection of the Cannes Film Festival exhibits a level of comedy different from that observed in the overall film population, we compare the observed statistic to a null distribution obtained via stratified bootstrapping from a reference dataset (TMDB). Strata are defined by year and main country of production. Each film is associated with a year y and a country c , and the Cannes selection is summarized by the number $n_{y,c}$ of films in each (y,c) cell. At each bootstrap iteration b , we construct a pseudo Cannes sample by randomly drawing, for each stratum (y,c) , exactly $n_{y,c}$ films from TMDB belonging to the same stratum. Because films selected at Cannes are often released after their production year, we allow a temporal tolerance by including films released in year y and $y+1$. For example, if in 2005 the Cannes selection includes 10 French films and 3 Brazilian films, we randomly draw 10 French films released in 2005 or 2006 and 3 Brazilian films released in 2005 or 2006 from TMDB. Sampling is performed without replacement when the TMDB pool is sufficiently large, and with replacement otherwise, in order to preserve stratum sizes.

To handle rare strata that are present in Cannes but absent from TMDB, we apply a hierarchical fallback strategy. First, if the exact (y,c) stratum is empty in TMDB, we expand the temporal window to $[y-2, y+3]$ while keeping the country constraint. If this expanded pool is still empty, we relax the country constraint and draw from TMDB films released in year y only. After aggregating all draws across strata, we compute the statistic of interest T_b on the pseudo sample. Repeating this procedure B times yields an empirical distribution of T_b representing the null hypothesis that the Cannes selection is a conditionally random sample given the year by country structure. Statistical significance is assessed by comparing the observed statistic T_{Cannes} to the bootstrap distribution using percentile confidence intervals and a two-sided empirical p value. A significant p value indicates that the observed level of comedy is too extreme to be attributed to random sampling alone.

We apply an almost identical procedure to the Cesar dataset. The main difference is that stratification is performed solely by year, since all films in both the Cesar selection and the corresponding TMDB comparison set are French. Regarding the temporal window, we restrict sampling to films released in the year preceding the award ceremony, in line with the official eligibility rules stating that eligible films must have been released the year before the ceremony. The same hierarchical fallback mechanism as described above is used to handle years that are underrepresented or absent in TMDB. By proceeding in this way, we approximate as closely as possible the institutional distributions of Cannes and the Cesar Awards and their respective eligibility criteria, while deliberately excluding genre information from the stratification, as genre is precisely the variable of interest in our analysis.

Each film is assigned an ordinal comedy score taking values 0, 1, or 2, where 0 indicates the absence of comedy, 1 corresponds to a mixed comedy, and 2 to a pure comedy. This score is derived from genre labels, based on the presence of the "comedy" tag and the possible co-occurrence of genres such as drama, romantic, action, or horror. The primary statistic of interest is the mean of this score within a given sample, which we interpret as the average intensity of comedy.

To assess how closely the bootstrap sampling scheme matches the target stratification, we report the frequency with which each sampling mode is used (exact match, country-window fallback, and year-only fallback). Because counting by strata may overrepresent cells with small sample sizes, we also report frequencies weighted by ny,c , that is, the proportion of films actually drawn under each sampling mode.

Results

The observed mean comedy score for Cannes selections is far lower than the value expected under the bootstrap null distribution ($\mu_1 = 0.241$ and $\mu_2 = 0.561$, see **Figure 4.k**), lies well below the 95 percent bootstrap confidence interval $[0.526, 0.596]$, and is therefore highly statistically significant ($p < 0.01$), indicating a strong underrepresentation of comedy in the Festival de Cannes nominees relative to the TMDB comparison baseline.

The observed mean comedy score for César nominees (0.510) is substantially lower than the value expected under the bootstrap null distribution ($\mu_1 = 0.510$ and $\mu_2 = 0.665$, see **Figure 4.l**), falls well outside the 95 percent bootstrap confidence interval $[0.575, 0.757]$, and is therefore statistically significant ($p < 0.01$) as well, indicating a marked underrepresentation of comedy in the César Award nominees relative to the TMDB comparison baseline.

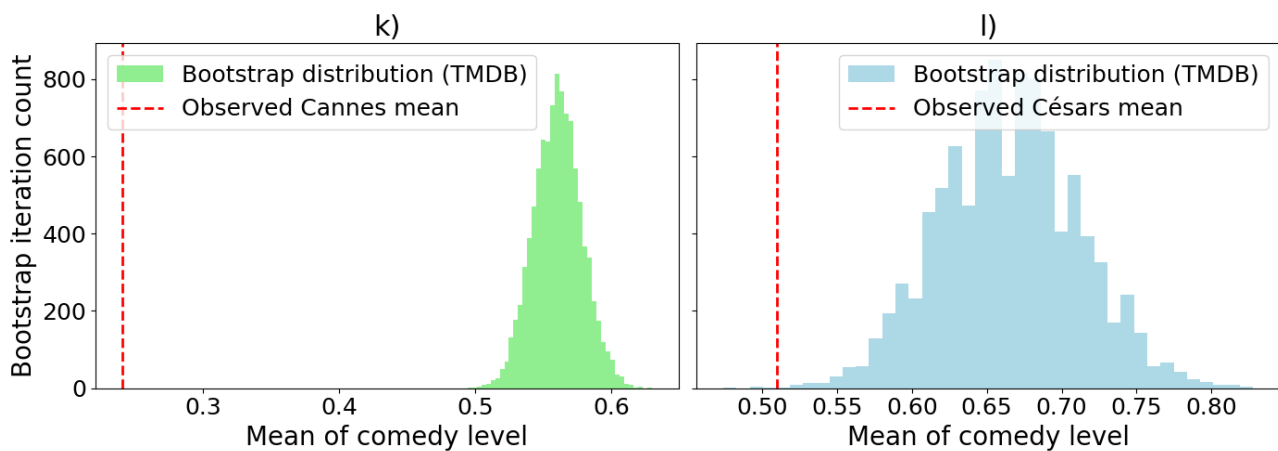


Figure 4 - Stratified bootstrap distributions of the mean comedy score for (k) the Cannes Film Festival selection and (l) the César Award selection, compared against their respective TMDB bootstrap baselines.

Conclusion

Although the winners of the top awards at the two major French film institutions do not exhibit a significantly lower level of comedy than the nominees, it could reasonably be argued that the number of observations is too small to reach statistical significance. The bias therefore lies elsewhere: nominees themselves display a significantly lower level of comedy than films that meet the eligibility criteria for selection.

That said, it is important to note that the TMDB dataset contains fewer films for the earliest years of these festivals than for later periods, which may raise concerns about the reliability of the results. One could argue that early years include only a limited number of films, likely skewed toward those that performed best commercially and, therefore, potentially toward comedies, which tend to perform better at the box office, as shown above.

Nevertheless, the consistency and robustness of the results provide meaningful insight into the selection biases exercised by juries in this type of competition, whether consciously or unconsciously.