A Data Scientist Analyzes American Movies and What They Say About Our Culture

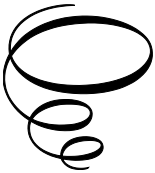
A Data Scientist Analyzes American Movies and What They Say About Our Culture:

Empire of Light

By

R. Grant Steen

Cambridge Scholars Publishing



A Data Scientist Analyzes American Movies and What They Say About Our Culture: Empire of Light

By R. Grant Steen

This book first published 2026

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-0364-6234-5
ISBN (Ebook): 978-1-0364-6235-2

The old world is dying, and the new world struggles to be born; Now is the time of monsters.
—Antonio Gramsci, Italian Marxist philosopher, written from a Fascist prison in the 1930's

Truth is so obscure in these times, and falsehood so established, that, unless we love the truth, we cannot know it.
—Blaise Pascal, in *Thoughts, Letters, and Minor Works, 1670*

Critics say that America is a lie because its reality falls so short of its ideals. They are wrong. America is not a lie. It is a disappointment. But it can only be a disappointment because it is also a hope.
—Samuel Huntington in *American Politics: The Promise of Disharmony, 1981*

TABLE OF CONTENTS

CHAPTER 1

MOVIES AND AMERICAN CULTURE

There's an odd thing about movies. If you love a movie, it's as if you saw it last week or last month, even if you saw it years ago. Some images remain as fresh in your mind as an artist's first stroke on an empty canvas, alive with possibility and promise, part of a story as unexpected as tomorrow's news. A movie can occupy a timeless place in your memory, beside your first crush or the first time you rode a bicycle. But, in truth, most movies are not remembered well or at all. What separates a timeless movie from a forgettable movie?

Artistry!

Is it legitimate to consider movies an art form, in the same sense as a painting or sculpture? To answer this question, we need first to define art, which has confounded critics and thinkers for years. I don't suppose I know better; all I can do is offer a definition of art that works for me.

Art certainly has an element of craft to it; setting aside non-representational art for the moment, since most movies are representational, a poorly crafted painting can leave a naive viewer wondering what they're looking at. And the fastest way to pull a viewer out of a picture, whether a painting or a movie, is to have some clanging artifact of poor craft; an errant brush stroke or an in-frame boom mike. Excellent craft can compensate for a banal subject or a sentimental viewpoint, but craft is not enough.

Art must also offer a fresh perspective, a new way of seeing something that is otherwise familiar. This is why Limoge china is not art, no matter how beautiful the piece; each piece is like every other of its type, so no new perspective is offered. The first singular piece of Limoge china was art; all that has followed is craft.

Art must also be accessible; there's little point to a work of art that is only understood by the artist or his mother. This criterion can be tricky, because viewers are loath to admit that they don't understand a piece, especially if

the craft is good. It can be somewhat traumatizing to admit that you don't comprehend the idea that animates a work, which may explain some works firmly in the canon.

The place of beauty in art is contentious. Certainly, Goya's painting of *The Third of May, 1808, in Madrid* is not beautiful, in the sense of being "pretty". Well balanced, evocative, colorful, dramatic, and yes, beautiful, in its intensity, theatricality, and tragedy; but not pretty. On the other hand, some works seem accessible because they are pretty—for example, David Hockney's watercolors—and it's easy to overlook that they are, as well, a new way of seeing.

There is also an emotional element to artistry; not the easy sentimentality that is the mark of a work trying too hard to be liked, but a deeper emotion, be it awe or reverence or horror or insight or even anger. Goya's *Third of May* packs a punch, even to people who know nothing of the history behind it. The emotions are visceral and accessible, partly because of the craft and partly because of the drama of the moment portrayed.

But the single feature that, to me, identifies a singular work of art is that it pulls you into itself, in a way that can't be explained. The first time I saw Picasso's *Guernica* at the Museum of Modern Art in New York, sometime prior to 1981, I was absolutely transfixed, frozen mid-step as I entered the room and first saw the huge work. I don't know how long I stood there, because I had fallen into the painting, gotten swept up in it, ravished by it, and made never the same. To this day, I have not forgotten that swept-away instant, though I am not ordinarily a fan of Picasso.

Are movies art? Of course, because you can get lost in them, can live a different life through them, can see things you've never seen before, feel things you're rarely felt before, know things you were never taught, be changed irrevocably by them and forget, if only for a moment, the world in which you live.

However, movies are essentially a corporate art form, different from a painting or a sculpture. Most artistic enterprises are undertaken by artists who are working alone; painting, sculpture, and photography are all lonely efforts. But movies are made more or less by committee; there are the actors who each contribute something unique and personal, the technical people who may say what is and isn't possible and, of course, no matter the power of the director, the money people. The only exception to the corporate nature of movies are experimental films, such as those made by

Stan Brakhage, an artist who worked alone, without actors, and often painted directly on the celluloid.

Most movies are made by large groups of people collaborating, a form of culture by committee. A movie is a director's output, but that output is constrained, since the director is limited in expression by the actors and by the effects that the technical crew can achieve, and the director answers to corporate money people. Directors can have a great deal of power, but a big flop can sink a studio, so directors are usually closely reined by the money people, who have ultimate control over the budget. The director is limited in what she can achieve because, without a great deal of money, a movie will not get made.

A studio greenlighting a movie is like a museum or gallery acquiring a piece of art, in the sense that the creative part of the enterprise will take place apart from direct corporate oversight; usually the corporation does not see the dailies. And, of course, both movie studios and museums are corporate entities, supervised by groups of people, who make compromises to appeal to their conception of what the public wants.

Because movies reflect the decisions of a group of people, movies are more likely to reflect culture at large than are most other art forms. The primary act of creation does not reside with the corporation, of course, but the corporation is choosing among possible alternatives. Art-by-committee can likely discern current culture more accurately than can an individual artist—even if that discernment is somehow subconscious—because the committees are concerned with making a product that is unique and novel, but not too idiosyncratic to succeed in the marketplace.

Movies are made by a committee whose main goal is to produce a movie that maximizes the potential audience. This makes it very surprising that Hollywood has a long history of flops, movies that do not earn out the money spent in making them. Between 2010 and 2020, Hollywood released an average of 738 movies a year (Statista, 2025. Accessed March, 2025), with quite a bit of variation year-to-year, and most of those movies were seen by relatively few people and didn't have a lot of impact. While some movies seemed destined to be a hit from the moment of conception—think *Titanic*, for example—most films are a middling success, even if a big audience was perceived. This means, of course, that even corporate entities are not all that good at predicting success for a movie. To complicate matters, every so often an unheralded movie hits it

big. If Hollywood could plan every movie to be a big hit, they would certainly do so.

There is a haunting history of surprise flops from movie studios. The biggest flop, in terms of money lost, within the range of this survey is probably *John Carter,* which was released in 2012 and lost over $200 million (Chmielewski, 2012). However, the most famous flop is probably *Heaven's Gate,* an epic Western from 1980. The director, Michael Cimino, was coming off a big success with *The Deer Hunter* and a smaller success with the well-crafted small-budget picture, *Thunderbolt and Lightfoot*. But somehow, *Heaven's Gate* flopped so grandiosely that, for a while, it became a punchline.

And then there are the initial flops that weather into sleeper hits; *The Rocky Horror Picture Show* did nothing at the box office for the first few months, then it slowly turned into a hit that plays still to a devoted audience that often yells key dialogue back at the screen.

There are also surprise hits, movies made with a small budget that vastly out-earn whatever was perceived as their box office potential. *Halloween* was made for $325,000 and earned $70 million, even though the picture had almost no advertising budget (Higgins, 2018). *Halloween,* by itself, started a movie franchise and revived the horror genre. Similarly, *Easy Rider* was made for about $340,00 but earned $50 million (American Film Institute, 2024) and came to represent a generation. Probably the movie that has the biggest ratio of profit-to-cost is *The Blair Witch Project*, which was shot in 8 days on a budget of $60,000, and became an astonishing critical and commercial hit, grossing almost $250 million (Deb, 2024). The profit-to-cost ratio for *Blair* easily exceeded the previous record-holder, which was *Mad Max.*

What accounts for the fact that studios seem unable to predict winners and losers? Could it be that movies interact with culture in a way that hasn't yet been characterized?

We will examine the relationship between a movie and the culture in which it lands. At the heart of this effort is a set of related questions: What makes some movies resonate?; How do movies capture the moment and become enormously popular?; What accounts for box office success? And what does box office success say about the culture into which the movie was released?

Sources

American Film Institute. “AFI Catalog of Feature Films. The First Hundred Years: 1893-1993”: AFI Catalog. Accessed April, 2024.

Chmielewski, Dawn C. 20 March 2012. “Disney expects $200-million loss on 'John Carter'“. *Los Angeles Times*. Retrieved April, 2024.

Deb, Sopan. 22 April 2024. “Blair Witch Project’ Actors Push for Retroactive Royalties”, *New York Times*. Accessed April, 2024.

Higgins, Bill. 7 September 2018. “Toronto Flashback: Millions Showed Up for ‘Halloween’s’ Scares in 1978” *Hollywood Reporter*. Accessed April, 2024.

Statista website. “Number of movies released in the United States and Canada from 2000 to 2024”: https://www.statista.com/statistics/187122/movie-releases-in-north-america-since-2001/. Accessed March, 2025.

CHAPTER 2

CAN WE DETERMINE IF A MOVIE HAS CULTURAL IMPACT?

America changed dramatically over the 80 years between 1940 and 2020. Demographically, we went from a country of 132 million people to one of 331 million people (United States Census Bureau. 1940 Census of Population). In 1940, 43% of Americans were rural, 90% were white, and only 5% of Americans over age 24 had a 4-year college degree (United States Census Bureau. 1940 Census of Population). In 2020, 14% of Americans were rural, 59% were white, and 34% over age 24 had a 4-year college degree (United States Census Bureau. Quick facts, 2020). Culturally, the differences between 1940 and 2020 seem even larger, so it is no surprise that the movies we watch have changed in nature.

We believe that movies reflect—perhaps even shape—the culture in which they land. Movies that became very popular in their particular era must have been closer to the *zeitgeist*, the spirit of that era. By examining the movies that were most popular, we might therefore be able to garner insights into that *zeitgeist*. We propose that movies that are box office hits make money specifically because they resonate with the culture in which they land. But, before we can rigorously test this idea, we must deal with a related issue. How do we objectively determine which movies resonated?

If the best movie ever made was playing at a nearby theater, we wouldn't know to see it unless we heard about it first, either by studio marketing or word of mouth. Studio marketing only succeeds if it can put a movie top-of-mind, which is hard to do. Word of mouth is usually seen as a better form of advertisement, since the recommendation might come from a trusted friend. Word-of-mouth advertising is a result of two things in relation to the viewer: enthusiasm and topicality. If a friend recommends a movie about something in which you have no interest, it is unlikely to result in your action. But if an enthusiastic recommendation from someone you trust relates to something of inherent interest to you, it likely will result in action. Viewer topicality is an individual thing, of course, but if throngs of viewers find a new movie to be of topical interest and

recommend it to a friend who also finds it interesting, then the movie has resonated with the culture.

Cultural resonance is therefore related to the popularity of a movie and may even determine its popularity. But how can we measure the popularity of a movie?

An obvious metric is the amount of money a movie makes at the box office. But there are evident problems with using box office as the only metric. How should the box office be quantified? Should we tally all the money that a movie ever made? That would seem to favor old movies over new, since the old movie had more time to accrue earnings. Should we tally only the money made in the first 12 months after release? That would discriminate against movies that took time to find their feet. It is also probably impossible to tally only the first 12 months after release with older movies, because bookkeeping records were not as rigorously kept in Hollywood's early days (Sackett. 1996). Should we tally the money made only in the calendar year of release? That would discriminate against pictures released late in the year, when some movies perceived to be strong Oscar contenders are released.

As an example of the subtleties of the box office problem, consider one of the highest grossing movies of all time, *Gone With the Wind*. Our analysis starts in 1940, the year when this movie made most of its money after first release. But the movie was actually released in December of 1939, so it does not even appear in our assessment. Also, when the movie was first released, ticket prices were far lower than they are today, so the take in dollars was rather small, without adjusting for inflation. While adjustment for inflation seems straightforward, one can always quibble with the results. Finally, *Gone With the Wind* was re-released many times. The first run was so successful that the studio chose to release it again in 1942, 1947, 1954, 1961, 1967, 1971, 1974, 1989, 1998, and, in the United Kingdom, in 2014 and 2024 (Wikipedia. *Gone With the Wind*. Accessed April, 2024). It seems discriminatory to exclude these later earnings, though most movies are never re-released.

There are further problems with using box office (BO) earnings as the sole measure of cultural resonance. Should we count BO earnings only in the United States, or should we also include earnings abroad? For the purposes of this analysis, insofar as possible, earnings in the United States will be considered in isolation. This is important because hefty earnings abroad are a recent phenomenon. In the past, earnings abroad were icing on the

cake; now they are factored in from the start, with some movies counted on to have more earnings abroad than domestically.

A final problem with using BO earnings as the sole metric is that some movies that critics agree were bad made a lot of money anyway.

Another possible metric to characterize cultural resonance is how many Oscar nominations a movie received in its year of release. This should compensate for the fact that some bad movies make a lot of money; such a movie would presumably not receive many Oscar nods. For this metric to be really useful, we need to focus on Oscar nominations as well as Oscar wins, simply because there are more of them. But a win should be more valuable than a nomination since it's harder to get a win. Still, certain weak performances receive Oscar nominations and even awards, especially if an actor or director had a lifetime of prior work that had gone unrewarded.

Yet another possible metric is some measure of current audience acclaim. Such acclaim should count whether rewarded with an Oscar nomination or not. Audience acclaim is particularly sensitive to movies that were perhaps somewhat ahead of their time; such a movie might have gone unnoticed at the box office or by Oscar but been noticed by audiences well after the fact. If we're still talking about a movie 50 years after it was made, then it must be important.

How should we choose among the possible metrics of box office earnings, Oscar nods, and audience acclaim? Why is it necessary to choose? A good metric could incorporate all these disparate sources of information, since each of them can fail in different ways. By combining metrics, it may be that we can obtain a more robust measure of the cultural impact of a movie. This approach may seem difficult, but it is easier than it might seem. If an objective measure of cultural resonance was available, it would enable us to take a measured and objective approach to the interaction between movies and culture.

Our focus will be entirely on successful movies. This is a departure from prior film criticism, in that virtually all critical film writing is based on anecdotes about individual movies that were well received by critics. In contrast, we will focus on a large cross-section of movies that were popular, with the idea that popularity is a surrogate measure of which movies resonate with the culture. Rather than focusing on a few movies anecdotally, we will focus on a large sample of movies, since this is more

likely to represent accurately the mass of humanity that is America. This large sample of movies will be analyzed using a scientific approach.

The scientific approach is based on data, or discrete observations often summarized in the form of measurements. Data is not the plural form of anecdote; it is quite different. Anecdotes are chosen to convey a point, to illustrate a claim a critic is trying to make, or to bolster a broad generalization. Data are objective facts, systematically gathered, that may or may not confirm the idea that is being investigated by the scientific approach.

What do we mean by a scientific approach? Science is a way of knowing, a method to approach the truth. It is not a field of learning; it is a process whereby we learn. If a person uses that process, she is acting as a scientist. While the process is difficult to learn, many people use it without being specifically taught.

The scientific method is based on a process of formulating a hypothesis, then testing that hypothesis with data. What is a hypothesis? Science is first and always based upon careful observation, the orderly accumulation of data. These observations relate to a set of facts that are initially not known. Once enough data has been accumulated, observations may suggest a tentative explanation, a proposal as to what mechanism may explain those observations. This tentative explanation is a hypothesis, the key element of the scientific process. It is crucial to remember that a hypothesis is a *tentative* explanation; where scientists are most liable to go astray is in coming to believe their own hypothesis in the absence of strong evidence for it.

To an English major, the concept behind a hypothesis might feel more familiar if called a thesis. The difference is that a hypothesis is a testable statement that forms the basis of a study, while a thesis is a statement that forms the basis of an argument that a paper is intended to support. A hypothesis is meant to be disproven, while a thesis is meant to be proven.

One can never prove a hypothesis; one can only fail to disprove it (Kuhn, 1996). But, if a hypothesis has been subjected to rigorous testing without being disproven, then it is provisionally likely to be at least a closer approach to the truth. This may seem like airy-fairy semantics, but it is not. The true scientist maintains a sense of skepticism and intellectual humility, knowing full well that the world is a far more complex place than we are ever likely to understand in depth. We can illuminate small

parts of it by careful study, but we are unlikely to ever have a thorough understanding of even seemingly simple processes. This is especially true when it comes to the workings of the human mind; it is an open question whether the human mind can ever truly understand itself (Steen, 2010).

A first-rate scientist actively seeks to disprove their own hypothesis. Only by seeking to break the tentative "new rule" that is a hypothesis, can we begin to have confidence in that insight. Besides, it is far better to disprove your own hypothesis than to have it disproven for you by another scientist, who will not be shy about calling out your failed hypothesis in the literature. Better to be humble and work to disprove your own hypothesis.

A hypothesis is basically just an idea, and it's rather worthless until it's tested. Anyone can have an idea; only a good scientist tries to disprove that idea. Science requires that a hypothesis be rigorously tested with new measurements. The basic question becomes, does the hypothesis predict these new observations? If the hypothesis is not predictive, then it likely isn't true. If a hypothesis is predictive, then it might be true, but there may be a more convincing explanation for the same facts that we don't yet know.

Hypothesis testing is not unique to science. Hypothesis testing is what a car mechanic does when you describe a problem with your vintage car. The advantage that a car mechanic has over a scientist is that the car mechanic has probably already encountered whatever you may describe. For example, if an old car won't start and only makes a clicking noise when the key is turned, there are three common causes which can be stated as testable hypotheses (Firestone Complete Aute Care website. "Your car won't start and you hear a clicking noise". Accessed May, 2024):

- **The car battery is dead**. If the battery fails, the car may make a rapid clicking noise because the battery can't turn the starter motor hard enough to get it started. A mechanic can test this hypothesis by using an instrument to measure the amount of charge left in the battery. If the battery charge has fallen below the level necessary to power the starter motor, then the car can be started with jumper cables. Once jumper cables are disconnected, if the car keeps running, the car battery was the problem. The hypothesis can be provisionally accepted.
- **The alternator has failed**. If the alternator fails, it can't charge the battery or provide sufficient power to the spark plugs, so sometimes a failed alternator causes the battery to run down. A mechanic can

test a hypothesis that the alternator itself is the problem, by jump-starting the car, then removing the jumper cables. If the engine stalls without the jumper cables, the problem is a failed alternator, because the spark plugs aren't getting sufficient current to spark.

- **The starter motor is bad**. If the starter motor fails, the car won't start, even if the battery is fine. A mechanic can test the hypothesis of a bad starter motor by first eliminating the more common cause of a dead battery. Then, carefully listening for the sound made when the key is turned can be helpful; a single click is likely to be a bad starter motor. If one then listens under the hood, it may be possible to hear a sound made by the starter motor itself, which is diagnostic.

It is certainly true that scientists are not the only ones who test hypotheses; this is done by anyone who seeks to understand the cause of a problem, including physicians, electricians, historians, economists, and even new mothers with a crying infant. It is also true that if you don't test your hypotheses, you're not doing science. Hypotheses can be tested by gathering new observations, to determine if the new data fits the predictions made by the hypothesis. If the new measurements are not as predicted, the hypothesis is probably wrong. This emphasizes the central role of measurement in science. Without measurement, science cannot be done.

It's worth taking a moment to say a word about "why?". Science is very good at addressing questions such as what, when, where, and who, but science cannot deal with why. The stories meant to explain the "why" of things—why giraffes have long necks (*e.g.*, so they can reach leaves high on a tree) or why gazelles are fast (*e.g.*, to avoid lions) —are just-so stories. These stories are myths that might seem true, but one can never be certain. Science can't address "why" questions because the question goes to motive. Motive assumes an overriding intelligence behind occurrences. An unassailable answer to "Why?" is, "Because God planned it that way." This claim cannot be tested—and certainly not disproven—so the claim is not in the realm of science. A belief that cannot be tested is not a hypothesis; it is the definition of faith.

We propose to test hypotheses about American culture, using data derived from an analysis of American movies. Our approach will be to use analytics. But what is meant by analytics? This term is probably familiar to many readers because analytics has long been used to evaluate the success of a football team or the impact of an individual basketball player.

Analytics is simply insight resulting from a systematic analysis of large quantities of data. As we noted before, data are systematically gathered information or variables that may (or may not!) have meaning.

Analytics is a sophisticated analysis of data; it is not in itself science unless analyses are used to test a hypothesis. Analytics is a search for patterns in the sea of noise that is everyday life. Any patterns that emerge can potentially give insight into the mechanism generating an observation. If analytics are used to test a hypothesis, then analytics can become a science.

Can analytics be used to analyze art? We propose that art can be measured, in terms of content and impact. Therefore, analytics can be applied to movies, just as it can be applied to football and basketball.

Which movies should be analyzed? This is an important question, for several reasons. First, a long span of years should be evaluated because culture changes rather slowly. Second, a large number of movies should be analyzed, because clear patterns can only emerge with a large sample size. Finally, we need to analyze popular movies that are likely to have resonated with the culture in which they arrived. Hence, our analysis will start in 1940, which was a period of remarkable unity in America culture. The analysis will run until 2020, which was a period of remarkable fragmentation and chaos. Because we are proposing to analyze 80 years of movies, this means that relatively few movies can be analyzed for any given year.

We propose to analyze the 20 most popular movies for every year between 1940 and 2020. This will mean analyzing 1,600 movies in total. A reasonable guess is that there will be about 50 observations per movie, for a total of 80,000 data points (see 17,600 data points collated in Appendix A). We will, at all times, try to maintain a clear distinction between a data point and the interpretation given to that data point.

We should acknowledge at the outset that not all measurements will be equally reliable. The fact that there are some unreliable measurements is a difficult problem. For example, movie studios did a rather poor job of tracking earnings in the early years. Inflation has had a huge impact on ticket prices, so it's hard to compare between years; Box office earnings are an estimate, a surrogate for contemporary and later popularity; There is manifest unfairness in comparing old movies to new, because old movies had more time to accrue earnings, but ticket prices were much lower. Still,

science is used to dealing with measurements that are not totally reliable, and, with a large enough sample size, analytics is generally robust to shaky data.

What questions can we potentially answer with an analytic approach to movies? A detailed exploration of what questions can be answered will constitute the bulk of this book, but here we can outline the kinds of questions that can potentially be addressed.

What were the most common types of movies then and now? How have the types of movies changed? Does a change in subject matter say something about the issues that concerned us then or concern us now? Has any potential change in movie type been abrupt or gradual?

Do changes in movie subject precede or follow changes in society? How close is the temporal linkage, if any? Is it plausible that movies cause cultural change, or do they merely reflect it?

Do hits and flops through the years give insight into contemporary American culture? If a movie resonates, is it more likely to be ahead of the culture or to reflect recent cultural changes that have already happened? Is a movie that flops more likely to be ahead of its time, behind its time, or simply out of the cultural mainstream?

Perhaps most important of all, have Americans lost faith in their own government, in the entity meant to bind us together into one people? Is this loss of faith reflected in movies?

Sources

Firestone Complete Aute Care website: “Your car won’t start and you hear a clicking noise”. https://www.firestonecompleteautocare.com/blog/maintenance/car-wont-start-clicking-noise/. Accessed March, 2024.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. University of Chicago Press; 3rd edition, 1996.

Sackett, Susan. *The Hollywood Reporter Book of Box Office Hits*. Billboard Books, 1996.

Steen, R. Grant. *The Evolving Brain: The Known and the Unknown.* Prometheus Books, 2010.

United States Census Bureau. 1940 Census of Population, Volume 2. Summary:

https://www2.census.gov/library/publications/decennial/1940/population-volume-2/33973538v2p1ch2.pdf
United States Census Bureau. Quick facts, 2020: https://www.census.gov/quickfacts/fact/table/US/PST040223
Wikipedia: *Gone With the Wind* (film) https://en.wikipedia.org/wiki/Gone_with_the_Wind_(film). Accessed April, 2024.

CHAPTER 3

THE CULTURAL IMPACT SCORE (CIS)

Science is based on the rigorous testing of hypotheses. But without meaningful measurements, hypothesis-testing is impossible. And, without hypothesis testing, science is not possible. Fortunately, scientists are skilled at developing meaningful measurements. We will develop a Cultural Impact Score (CIS) to characterize the extent to which a movie interacts with the culture in which it lands. The CIS will be used to generate hypotheses that can then be objectively tested.

Science is accustomed to dealing with suboptimal or imprecise measurements. In the early stages of a science project, it is often not entirely clear what measurements should be gathered. Even after clarity is achieved as to what measurements will be most useful, it can still be hard to determine how best to make those measurements. This discovery process—learning how to make relevant measurements—must be logically developed at the outset, clearly described as it evolves within the constraints of the system, and consistently used after development. The scientist must be explicitly clear as to how all measurements were made and what assumptions underlie those measurements, so that other people can replicate the results, thereby further testing the hypothesis.

Our CIS measurement will be based on movie box office in the year of release, Oscar wins and nominations in the year following release, and a current measure of audience acclaim. We will describe in detail how each component measurement is obtained, so that our work can be replicated.

The most reliable source of box office information is *The Hollywood Reporter*, a Los Angeles-based trade magazine that focuses on the entertainment industry. Recently, *The Hollywood Reporter* ranked the 5 most successful films for each year, from 1939 to 1995 (Sackett. 1996). For movies outside of this span of years, and for movies below the top 5 for any given year, the reference of choice is IMDbPro (IMDb Pro. Accessed May, 2024), a database published by Amazon and used extensively by movie industry professionals. IMDbPro collates earnings

for every film released in a given year, so earnings can be ranked in that year, in a sub-section called “In-year releases”.

However, we will not use the actual dollar value of the box office (BO) in any calculation, since BO is not entirely reliable, because of the problems discussed in Chapter 2:

- **Reliable BO for the calendar year of release is unavailable for early films**. Bookkeeping was somewhat scattershot in Hollywood’s early days, and first-run BO is often inaccurate or contentious (Sackett. 1996).
- **BO numbers for the first 12 months after release are not available until quite recently**. Though this would probably be the best measure of timely success, this number was not collated at all until long after the early days of Hollywood.
- **BO numbers often include earnings to date for re-released movies**. This makes for an unfair comparison between movies that were re-released and those released only once.
- **BO numbers can include earnings abroad**. Earnings outside of the United States are not relevant to our discussion.
- **BO numbers from older movies must be adjusted for inflation**. Earnings for the top-grossing 1940 movie *Fantasia*, which is still famous, were about $23.3 million, while *Like A Boss*, the fourteenth most profitable movie of 2020, which hardly anyone remembers, were about the same at $22.2 million (2). In 1940, the average U.S. house price was under $3,000, according to the U.S. Census Bureau; by 2020, the average U.S. house price was $329,000 (U.S. Department of Housing and Urban Development. Table 8, New Single-Family Home Prices: 1964–Present. Accessed May, 2024), more than 100-fold higher than the 1940 house cost. Thus, inflation can play havoc with estimates of earnings.

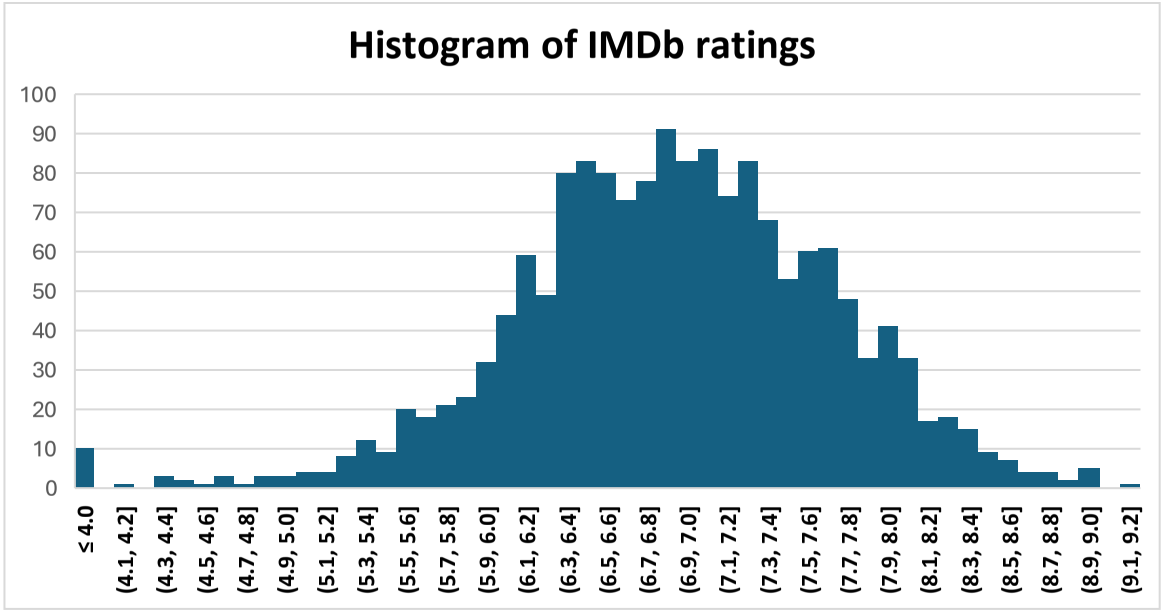
The actual dollar value of the BO is therefore considered a “guesstimate” and will not be used in any calculations. Instead, we will use the relative rank of earnings in the year of release as a measure of audience interest at the time. Relative rank is more likely to be correct in a given year and is robust to comparisons between years. The logic of using only the top 20 movies in a given year is that these movies would have played at virtually every neighborhood theater.

We will also use the metric of Oscar nominations and wins because these reflect esteem within the movie industry. We note that nominations are

more useful than wins, simply because there are more of them, but wins should be worth more than mere nominations. Furthermore, every Oscar nomination and win should count, not just those for best picture or best director. A win for best special effects is still a win, though most people outside of the industry are unlikely to remember it.

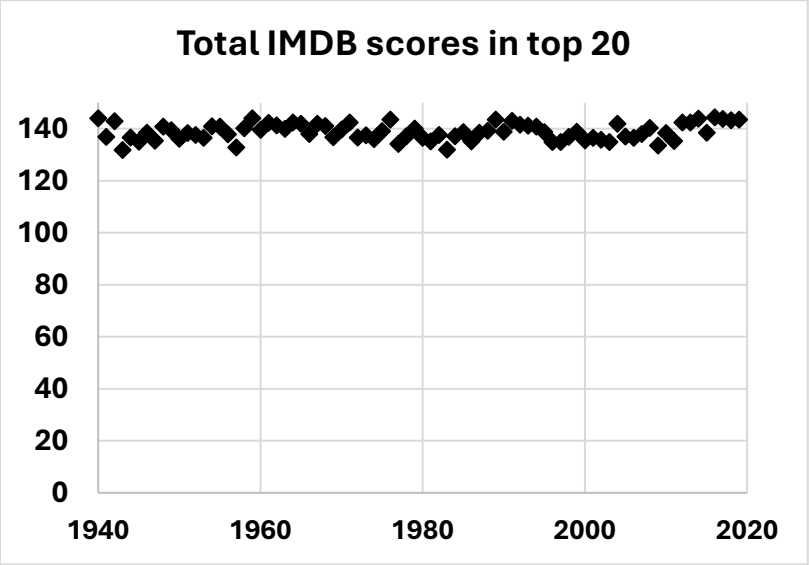
Finally, we will use a measure of current audience acclaim as an important metric. Several on-line databases track audience ratings, so this gives an opportunity to assess the quality of the ratings. The Internet Movie Database (IMDb), which is owned by Amazon, is one possible source, whereas Rotten Tomatoes (RT), at least partially owned by the movie studios, is another possible source. The correlation in audience rankings between IMDb and RT in 1940 was 80%, so the measures are similar, but not identical. RT is somewhat less complete than IMDb; the movie *Rebecca*, which was the fifth-ranked movie at the box office in 1940, was missing from the RT database when I checked in May, 2024. Furthermore, IMDb routinely has movie trailers available at the website, whereas RT does not.

IMDb scores for the 1,600 movies in our database tend to be clustered around an average value of 6.94 out of 10. When plotted as a histogram, IMDb scores look like this, forming the typical bell curve:


Histogram of IMDb ratings
100
90
80
70
60
50
40
30
20
10
0
≤ 4.0
(4.1, 4.2]
(4.3, 4.4]
(4.5, 4.6]
(4.7, 4.8]
(4.9, 5.0]
(5.1, 5.2]
(5.3, 5.4]
(5.5, 5.6]
(5.7, 5.8]
(5.9, 6.0]
(6.1, 6.2]
(6.3, 6.4]
(6.5, 6.6]
(6.7, 6.8]
(6.9, 7.0]
(7.1, 7.2]
(7.3, 7.4]
(7.5, 7.6]
(7.7, 7.8]
(7.9, 8.0]
(8.1, 8.2]
(8.3, 8.4]
(8.5, 8.6]
(8.7, 8.8]
(8.9, 9.0]
(9.1, 9.2]

There are few movies with a score less than about 5.0, among the top 20 movies ranked here, just as there are few movies with an IMDb score as high as 9.0. What this will mean, in practice, is that IMDb scores will not give the CIS a broad spread of values. The factor likely to contribute most to the spread of possible CIS values is box office ranking, which runs from 1 to 20.

An important question is whether IMDb rankings are stable over time. It is easy to imagine that newer movies might be more highly ranked than older movies, because older movies can seem dated. To assess this possibility, we summed all IMDb audience scores for the top 20 movies of each year from 1940 to 2020. When plotted, total IMDb scores in each year look like this:



What does this graph show?

- **IMDB rankings are remarkably stable through the years**
- **IMDB rankings appear to be a useful metric**

It is now possible to specify a metric that combines ranked BO (1 through 20), Oscar nominations (noms) and wins, and IMDb score in a measure that we will call the CIS:

Cultural Impact Score (CIS)

CIS = (21-BO rank) x (1 + [2 x {Oscar noms + Oscar wins + Best picture noms + Best picture wins + IMDb rating}])

The first term (21-BO rank) stipulates that a movie ranked number 1 at the box office is worth 20 points, while a movie ranked number 20 is worth only 1 point. CIS is then calculated as the box office rank of a movie in a given year times a complex product made up of the sum of Oscar nominations, Oscar wins, Best Picture nominations, Best Picture wins, and IMDb rating. To win a Best Picture Oscar, a movie obviously has to be nominated, so a win is worth more than a nomination. Best Picture nominations and wins count separately but also count as a part of total Oscar nominations and wins. Thus, a Best Picture win is worth a total of 8 points. In order that a movie with no Oscar nominations or wins does not wind up with an artificially low score, there is a term at the beginning of the parenthetical expression ("1") that means the movie's CIS score would be determined entirely by the movie ranking at the box office and the IMDb score.

At this point, we need to discuss statistics and how they will be used, at least briefly and using a minimum of math. What is the purpose of statistics in science? The simplest answer is statistics are used to test hypotheses. If you state a hypothesis and gather data to test that hypothesis, how do you know if the data are strong enough to confirm your hypothesis? Eyeballing the data is not good enough; even the best scientists are motivated to trust their intuition and believe their own data. Statistical tests are used to determine objectively whether the data disprove or fail to disprove the hypothesis.

A range of tests are used, and it is beyond our purpose here to discuss the various tests and how they are used. Suffice it to say that generations of mathematicians have devoted time and effort to defining how to test various types of hypotheses. Basically, when data are gathered, you plug the numbers into an appropriate test, and you get a p value. The p value is the probability that the results were obtained entirely by random chance, in the absence of the mechanism proposed. A p value can range from 0.00 up to 1.00, with the upper bound ("1.00") indicating that there is no relationship

whatsoever and the lower bound ("0.00") indicating that the relationship is so strong that it is effectively perfect. In 40 years of doing science, I've never seen either of these values; they are really just mathematical limits. Most commonly, you get a p value somewhere in the middle. By convention—and really, convention is the only reason—a p value less than or equal to ($\leq$) 0.05 is considered to be "significant". When $p = 0.05$, this means that the odds of obtaining the same data purely by chance—in the absence of a real relationship—are 5 in 100 (or about 1 in 20).

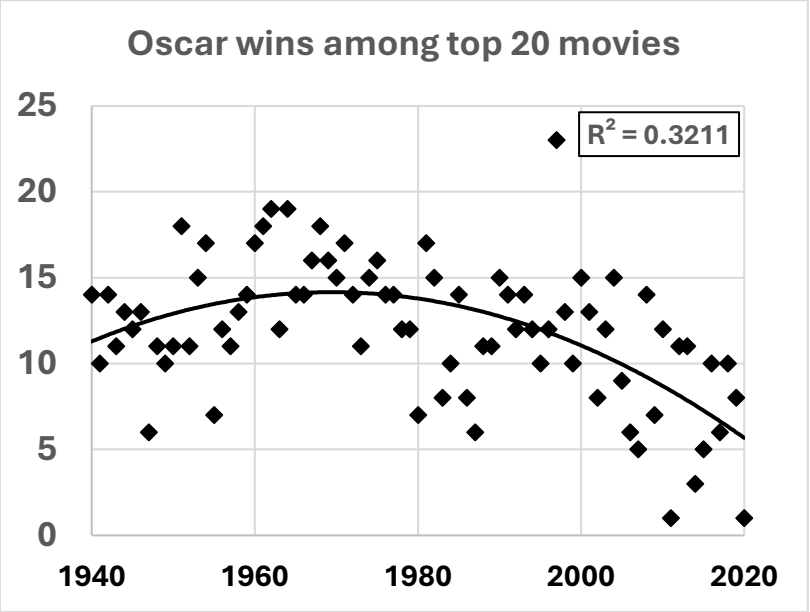
What does significance actually mean? If a relationship isn't significant, it isn't a relationship. If a change isn't significant, it isn't a change. A change cannot be "somewhat significant", just as a woman can't be "somewhat pregnant". A significant relationship between two variables means that one variable is potentially explained by the other variable.

What factors affect whether a relationship is significant? First, and most obviously, if there really isn't a relationship, you likely won't achieve significance when doing a statistical test. But, even if your proposed relationship is real, you may fail to achieve a statistically significant result if your sample size is too small. This is why we will evaluate 1,600 movies; this sample size is large enough that even a weak (but real) relationship will likely achieve significance.

Let's look at an example of how statistics are used. In collating the data, I had the impression that, among the top 20 movies in recent years, there were fewer Oscar wins than in the past. When I plotted total Oscar wins year-by-year, I got the graph below. The Excel spreadsheet I used enabled me to fit a line to these points that looked like a reasonable fit. A fitted line implies a correlation that is actually an explanatory hypothesis. The hypothesis is that there is a meaningful relationship between Oscar wins and year of release, and a statistical test of this hypothesis yielded an R^2 value of 0.3211. R^2 is generally interpreted as the proportion of variability in one value that is explained by variability in the other value. In this case, approximately 32% of the variability in Oscar wins is determined by the year of release, suggesting that Oscar sees less worth in the popular pictures of recent years. But is this result statistically significant?

What does the graph below show?

- **Oscars wins among top 20 pictures are declining**
- **Peak Oscar awards in a given year was in the 1960s**
- **The outlier year (1997) was the year of *Titanic***
- **Oscar now rewards movies that are less successful at the box office**



We used an on-line calculator (Social Science Statistics. P Value from Pearson (R) Calculator. Accessed May, 2024) to determine that, if R^2 = 0.3211 and the sample size is 80 years, then $p<0.00001$. This means there is one chance in 100,000 that these results are an accidental finding, and not the result of a true relationship. The initial impression, which was turned into a hypothesis and tested, has been confirmed or, more truthfully, not disproven. In fact, we are not likely to have gotten results this good merely by chance, given that the odds of an accidental finding are 1 in 100,000. Therefore, we can provisionally accept that there is a disconnect between the box office success of a movie and whether that