rotten_tomato_movie_review.csv

Variable Definitions

1. **movie_id**: Unique identifier for each movie (matches the id in rotten_tomatoes_movies.csv)
2. **reviewId**: Unique identifier for each critic review
3. **creationDate**: The date the review was published
4. **criticName**: The name of the critic who wrote the review
5. **isTopCritic**: A boolean value indicating if the critic is considered a top critic
6. **originalScore**: The score provided by the critic
7. **reviewState**: The status of the review (e.g., fresh, rotten)
8. **publicatioName**: The name of the publication where the review was published
9. **reviewText**: The name of the publication where the review was published
10. **scoreSentiment**: The sentiment of the critic's score (e.g., positive, negative, neutral)
11. **reviewUrl**: The URL of the original review on Rotten Tomatoes

Summary Statistics

	movie_id	reviewId	creationDate	criticName	isTopCritic	originalScore	reviewState	publicatioName	reviewText	scoreSentiment
count	1444963	1.444963e+06	1444963	1444963	1444963	1009745	1444963	1444963	1375738	1444963
unique	69263	NaN	8510	15510	2	1729	2	2707	1359771	2
top	parasite_2019	NaN	2000-01-01	Emanuel Levy	False	3/5	fresh	New York Times	Parental Content Review	POSITIVE

rotten_tomato_movies.csv

Variable Definitions

1. Movie_ id: Unique identifier for each movie (matches the id in rotten_tomatoes_movies.csv)
2. title: title of the movie
3. audienceScore: The average score given by regular viewers
4. tomatoMeter: The percentage of positive reviews from professional critics
5. rating: The movie's age-based classification (e.g., G, PG, PG-13, R)
6. ratingContent: Content leading to the rating classification
7. releaseDateTheater: The date the movie was released in theaters
8. releaseDateStreaming: The date the movie was released to streaming
9. runtimeMinutes: The duration of the movie in minutes
10. genre: The movie's genre(s)
11. originalLanguage: The original language of the movie
12. director: the movie's director
13. writer: the movie's writer(s)
14. boxOffice: The movie's total box office revenue
15. distributor: The company responsible for distributing the movie
16. soundMix: The audio format(s) used in the movie

Summary Stats

	movie_id	title	audienceScore	tomatoMeter	rating	ratingContents	releaseDateTheaters	releaseDateStreaming	runtimeMinutes	genre	originalLanguage	director	writer	boxOffice	distributor	soundMix
count	143258	142891	73248.000000	33877.000000	13991	13991	30773	79420	129431.000000	132175	129400	139041	90116	14743	23001	15917
unique	142052	126403	NaN	NaN	10	8353	12062	4726	NaN	2912	112	62206	67274	4863	3693	551
top	junction_48	The Stranger	NaN	NaN	R	['Language']	2018-09-14	2017-05-22	NaN	Drama	English	Unknown Director	Jing Wong	$1.1M	Paramount Pictures	Surround