

# Weekly Report

Mike Ma

September 12

## 1 Brief Summary

There is a Brief Summary of what I have done recently.

1. Having solved the title matching problem, and merge the trailer data from Youtube and movie data from IMDb into one table.

(This means we correspond specific movie details for each trailer and we could use trailer data to predict movie now.)

2. There are new problems after title matching, the movie data are incomplete, which means some trailer couldn't find corresponded movie data. Therefore, I am collecting the whole movies information from IMDb from 2011-2019, about 9000 movies per years. The software is still running, but I have already written the title-matching codes for this new movie data table. By tomorrow, I could achieve a much more complete merged table.

(I could collect everything you see in the first page, but I couldn't get access to the second page.)

[←](#)
[→](#)
[↺](#)
[🔒](#)
nytimes.com/reviews/movies

# Movie Reviews

Our film critics on blockbusters, independents and everything in between.

---

MPAA Rating

Genre

☐ NYT Critic's Pick

---

Sept. 12, 2019  
[Read Review](#)  
[Find tickets](#)

## Ms. Purple

NYT Critic's Pick | Drama | Directed by Justin Chon

A moody, downbeat drama that occasionally offers glimmers of hope, Justin Chon's film follows two Korean-American siblings struggling to reconnect while their father is dying.

By JEANNETTE CATSOULIS

---

Sept. 12, 2019  
[Read Review](#)  
[Find tickets](#)

## América

NYT Critic's Pick | Documentary | Directed by Erick Stoll, Chase Whiteside

Filmed in Mexico, this tender and subtle documentary follows the struggles of three young men to care for their beloved aging grandmother.

By GLENN KENNY

---

Sept. 12, 2019  
[Read Review](#)  
[Find tickets](#)

## Freaks

R | Drama, Sci-Fi, Thriller | Directed by Zach Lipovsky, Adam B. Stein

This debut feature, starring Emile Hirsch as the paranoid father of a cloistered girl, borrows heavily, and happily, from comic-book franchises.

By JEANNETTE CATSOULIS

---

Sept. 12, 2019  
[Read Review](#)  
[Find tickets](#)

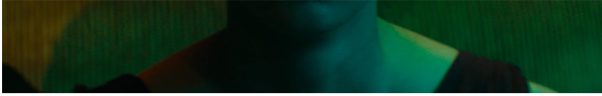
## Depraved

Drama, Horror, Thriller | Directed by Larry Fessenden

Larry Fessenden updates Mary Shelley's classic tale, "Frankenstein," producing possibly his most coherent and visually polished work yet.

By JEANNETTE CATSOULIS

nytimes.com/2019/09/12/movies/ms-purple-review.html



Tiffany Chu plays a conflicted young woman in Justin Chon's "Ms. Purple."  
Oscilloscope Laboratories

By **Jeannette Catsoulis**

Sept. 12, 2019, 7:00 a.m. ET

**Ms. Purple** NYT Critic's Pick | Directed by Justin Chon | Drama | 1h 27m

[FIND TICKETS](#) When you purchase a ticket for an independently reviewed film through our site, we earn an affiliate commission.

Following two deeply damaged siblings, each lacking a place in the world, Justin Chon's "[Ms. Purple](#)" seems named not for a character, but for a state of mind that's been a long time brewing.

Purple is also the color of the traditional South Korean dress obediently worn one evening by Kasie (Tiffany Chu), 23, at the insistence of her rich, entitled boyfriend (Tony Kim). But in the United States, where the film takes place, purple vividly signifies daring and defiance, independence and strength. That demands a personality to match, and Kasie is a woman controlled by the demands of men. To her fond-enough boyfriend, she's property, a sex partner and a compliant plus-one. To the boozed-up businessmen at the high-end karaoke bar where she works as a hostess, she's a body to be fondled and shared and sometimes drugged. And to her dying father (James Kang), lying comatose in her childhood home, she's a caregiver whose devotion is necessarily unrequited.

When the aide tending to her father abruptly leaves, Kasie is forced

3. Having eliminated the invalid data.  
(for example, there are some videos "New Best Movies Trailers top 10" rather than trailer for specific movie.)
4. Having extracted the series number for each trailer.
5. Having organized the data and stored them into google cloud.
6. Having tried collect movie reviews from New York Time. I could get pre-released movie professional reviews from New York Time. I now agree that this is a good data resource. Each movie has one professional review, and it is usually a long article. However, what I could collect now is the "brief review summary" of each movie. It is usually the first paragraph of the article.  
(Because to get the full review article New York Test would check "I'm not a robot")

## 2 Method of title matching

1. Export the trailer data table and movie data table from MySQL into Json.
2. Create list for each field.
3. Use re.match() to match.
4. Check the format of unmatched part, if it included subtitle or series number, the matched part would not be considered.
5. store all the fields into one database.

## 3 Next Step

1. Collect the merged data as complete as possible.
2. Collect the remarks from New York Time.
3. Select the trailer comments in a fixed period before publish time.
4. Comments Sentiments Analysis.

## 4 Current Problems

1. I can't get the whole comment article in New York Time. Each movie has one professional review, and it is usually a long article. However, what I could collect now is the "brief review summary" of each movie. It is usually the first paragraph of the article.
2. We need to decide how long the fixed period should be.(Select the trailer comments in a fixed period before publish time. Like all comments one year before the trailer was published)
3. After the sentiments analysis, what we could only get is just a number of "like", "dislike", and "neutral". However, the amount of sentiment analysed trailers are much smaller than the current trailers(All current trailers have "like" and "dislike" count but not valid because of trailer published at the different time.).