Film Social Networks: Friend Effects With Individual-Level Panel Data From Letterboxd.com

Johnny Ma
University of Chicago

June 9, 2017

Introduction

Films are taste-based goods that are important to society.

- American Hollywood Industry and Cultural Significance of Film
- Artistic Mass Media sold to global consumers
- Subjective quality, dependent on many latent variables
- Netflix and FilmStruck have made diverse film content easily accessible for all users
- Who doesn't watch movies?

Motivation

What makes a movie flop? Why was *The Avengers* wildly popular but *Superman v Batman* a horrible failure?

More recently, did good reviews boost *Wonder Women's* sales and ratings more than in a world with social learning?

Are there some critics and central figures who define taste? (Roger Ebert/Anthony Fantano Effect).

Can consider anchoring/behavioral explanations, but regardless observations of peers and experts leads to changes in behavior and taste.

What Film To Watch Tonight?

Given that film is subjective art, difficult to understand why people watch what they do.

- Convenience (what is playing in theaters)
- Personal Taste Formation (Becker & Murphy)
- What Others are Watching: Aggregate and Local

Can we look at the effect of film and group averages of ratings on watching and rating behavior?

Moretti (2008)

Moretti (2008) sets up a model of "expected appeal" and information from peers and tests using aggregate sale data.

- Sign of realization over expected quality diverges sales
- Some notion of priors, some notion of type of shock (weather no effect)
- Positive shocks are strong for those in large social networks
- Can estimate some social multiplier

Overall an interesting model. We will borrow the idea of "surprise" in reducing risk. Not sure how reliable aggregate sales data can ever be.

Social Intuition

People often watch movies in groups, so collective decision making is important.

People actively consume reviews (Rotten Tomatoes), some critics are "taste-makers". Linear-in-means type model.

People ask friends for film suggestions based on their taste and their friends taste.

Recommendation Systems use correlations between users to create suggestions (Bell & Koren)

People read **laymen reviews**, more and more so these days. This, in a social network context, is our focus.

The Idea

People have social networks of film-watchers whose behavior and preferences are observable.

Individuals gather data about other individual's taste by their watching behavior and their rating behavior.

When choosing which movie to watch, they rely on the "surprise" of their friends to determine the lowest risk/highest reward film, conditional on their own taste and the characteristics of the films they are considering.

Model is largely similar to Moretti.

The Model

$$Pr(Watch_{ij}) = \alpha_i + \beta_1 * s(good)_{k \in i,j} + \beta_2 * taste_i + \beta_3 * film_j + \epsilon_i$$

j films, i individuals. Each i individual is in a network of k 'friends'. This is a group fixed effects type regression. Taken a bit from Conley and Udry (2010).

Taste is defined as the movies people have watched before (binary over films), film is a dummy for each film (J x J Identity). Other ideas are user ratings (Recommendation System Latent models) and film characteristics (director, language, year, actor, genre, etc.).

Identification Problems

The main problem: Endogenous Network Formation.

Along with Reflection Problem, one can easily say that any positive β_1 is not because of behavior but because people who watch similar movies tend to be friends. Indeed, highly correlated (not shown). Attempt to attenuate this problem with user and film level controls, but valid critique.

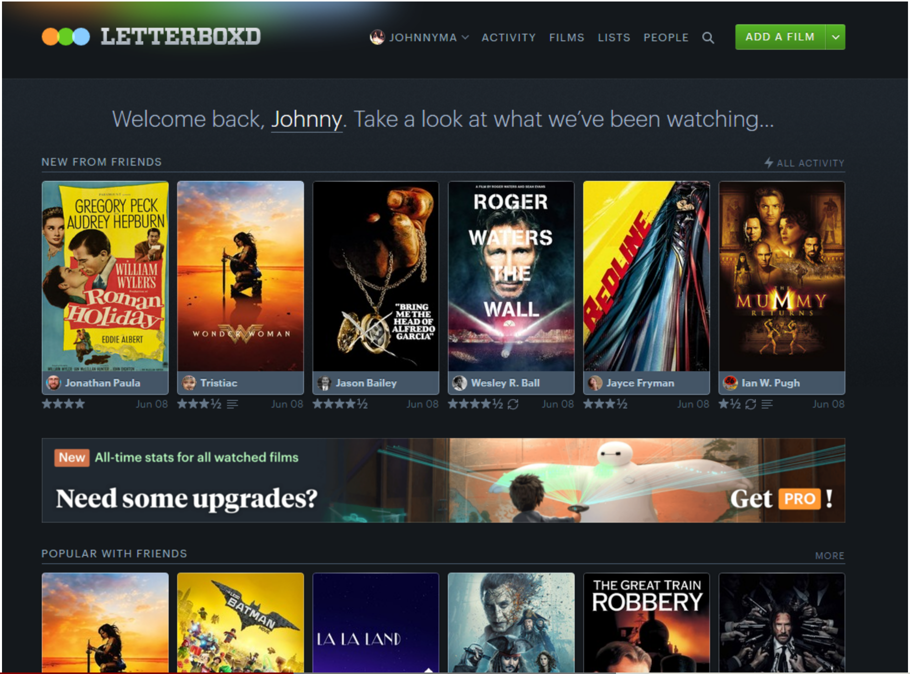
Bramoulle, Djebari, FortÃŋn (2009) develop a friend-of-friend-but-not-my-friend IV approach to identify peer effects in social networks, may consider for future work.

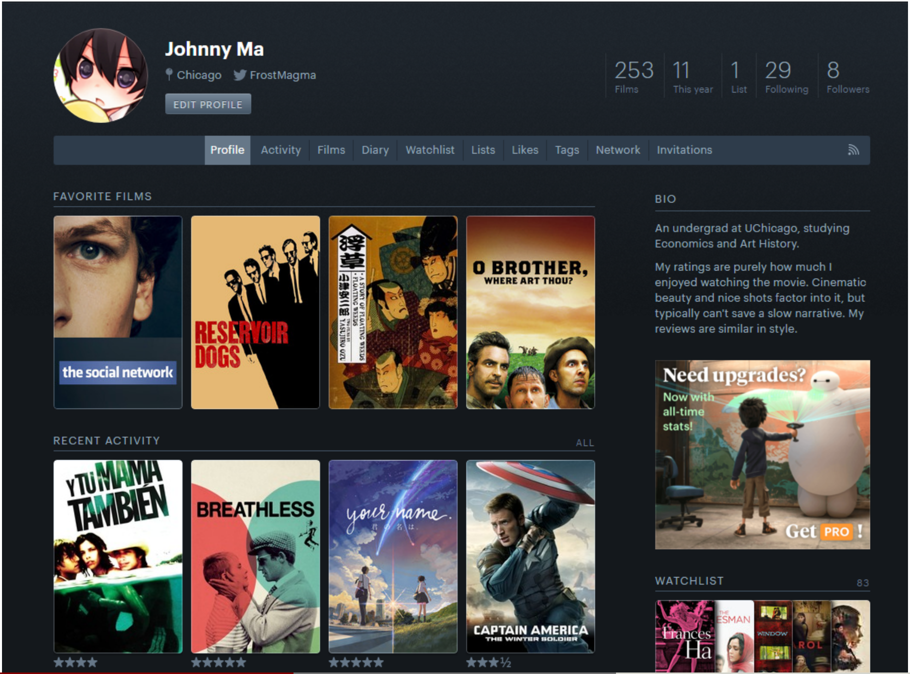
Regardless, want to see if friends characteristics or behavior have anything to do with user behavior.

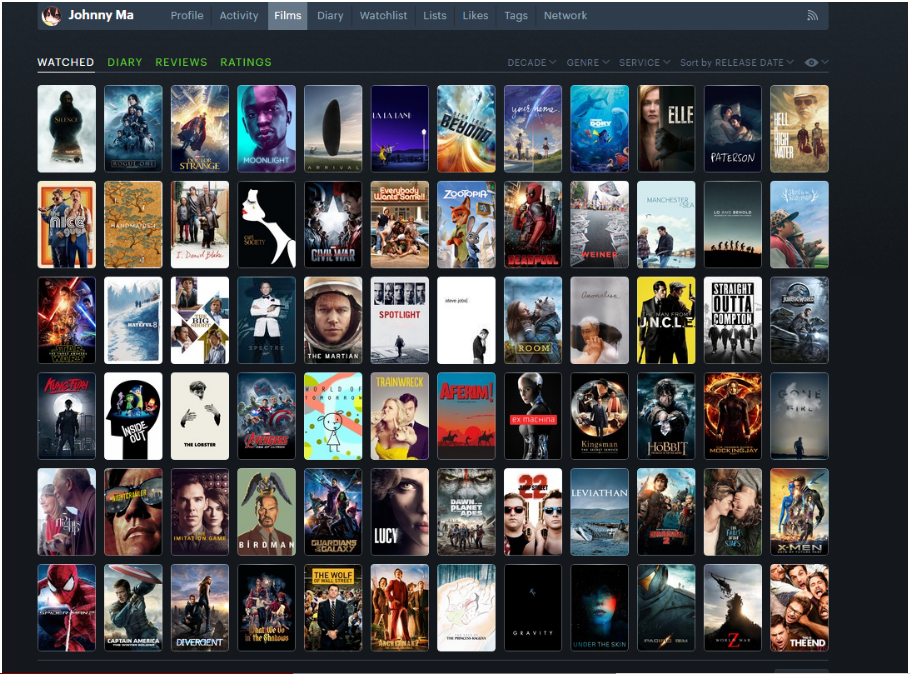
Micro-level data from *Letterboxd.com*, a film diary/social network website.

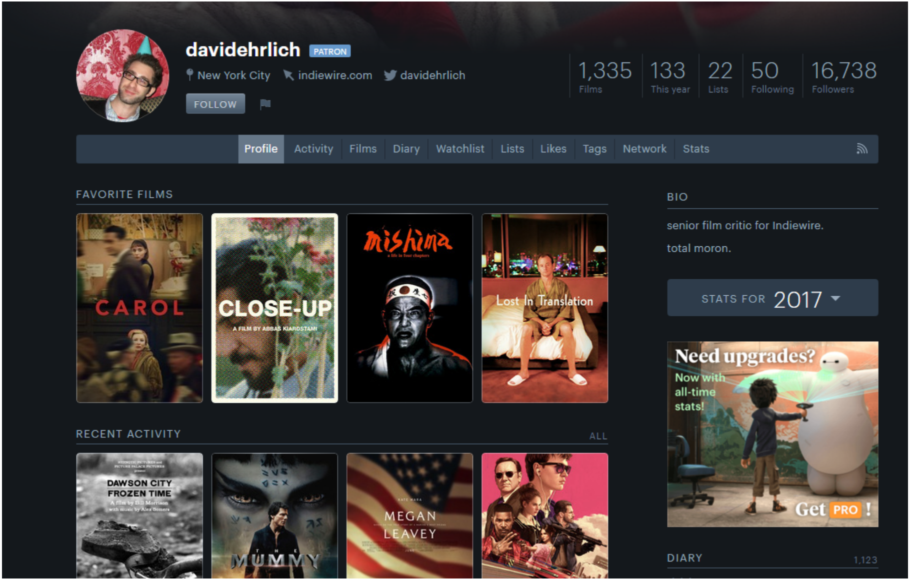
Small but growing community of film fanatics, ranging from CEO of IndieWire to Professional Bloggers to College Students (me). Average films seen is around 1,200 (that is a LOT, 1800 hours).

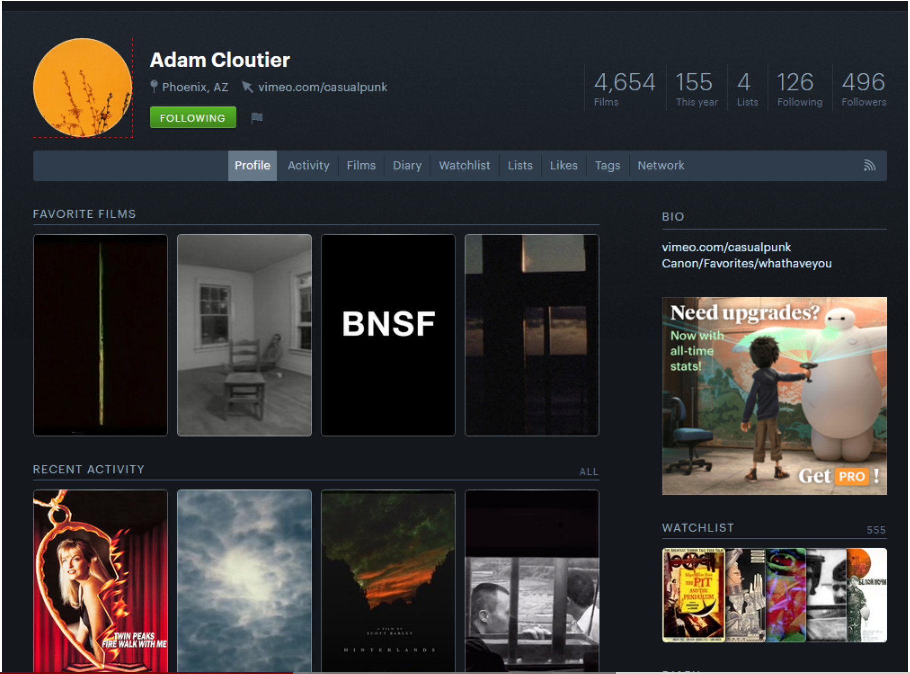
Typically used as film diary and film information (IMDB), though user reviews become popular and networks do form.

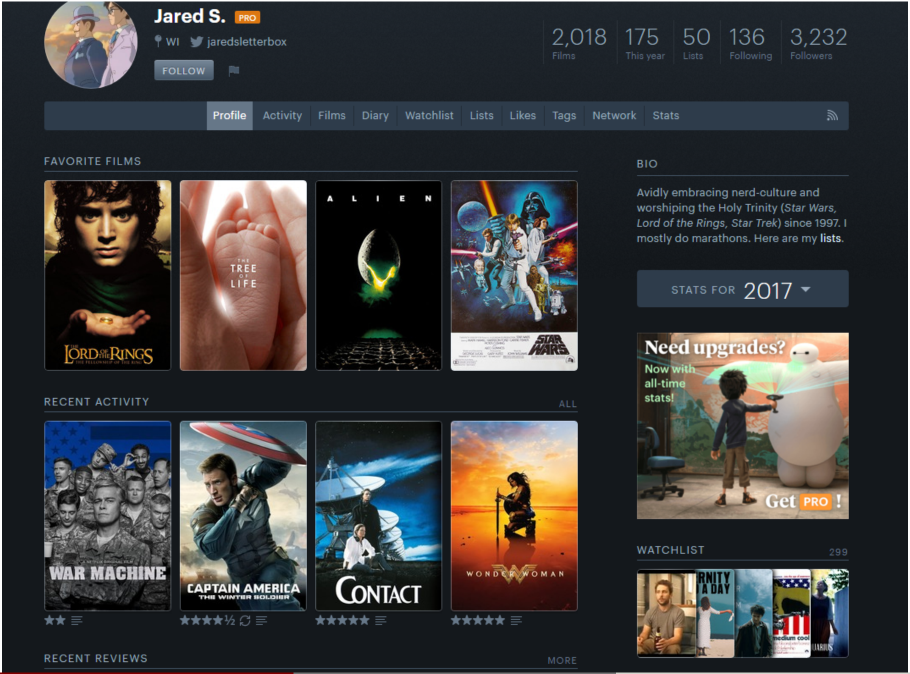

LETTERBOXD
JOHNNYMA
ACTIVITY
FILMS
LISTS
PEOPLE
ADD A FILM
Welcome back, Johnny. Take a look at what we've been watching...
NEW FROM FRIENDS
ALL ACTIVITY
GREGORY PECK
AUDREY HEPBURN
WILLIAM WYLER'S
Roman Holiday
EDDIE ALBERT
WONDER WOMAN
"BRING ME THE HEAD OF ALFREDO GARCIA"
ROGER WATERS THE WALL
REDLINE
THE MUMMY RETURNS
Jonathan Paula
Jun 08
Tristiac
Jun 08
Jason Bailey
Jun 08
Wesley R. Ball
Jun 08
Jayce Fryman
Jun 08
Ian W. Pugh
Jun 08
New
All-time stats for all watched films
Need some upgrades?
Get PRO !
POPULAR WITH FRIENDS
MORE
BATMAN
LA LA LAND
THE GREAT TRAIN ROBBERY

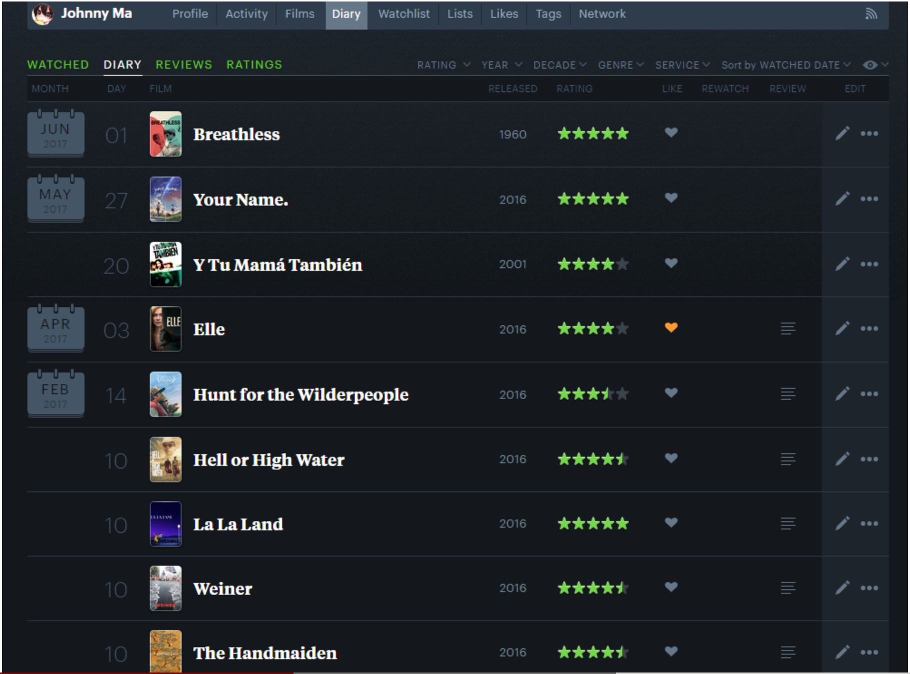

Johnny Ma
Chicago
FrostMagma
EDIT PROFILE
253 Films
11 This year
1 List
29 Following
8 Followers
Profile
Activity
Films
Diary
Watchlist
Lists
Likes
Tags
Network
Invitations
FAVORITE FILMS
the social network
RESERVOIR DOGS
O BROTHER, WHERE ART THOU?
BIO
An undergrad at UChicago, studying Economics and Art History.
My ratings are purely how much I enjoyed watching the movie. Cinematic beauty and nice shots factor into it, but typically can't save a slow narrative. My reviews are similar in style.
Need upgrades?
Now with all-time stats!
Get PRO !
RECENT ACTIVITY
ALL
Y TU MAMÁ TAMBIÉN
BREATHLESS
your name.
CAPTAIN AMERICA
THE WINTER SOLDIER
WATCHLIST
83
Frances Ha

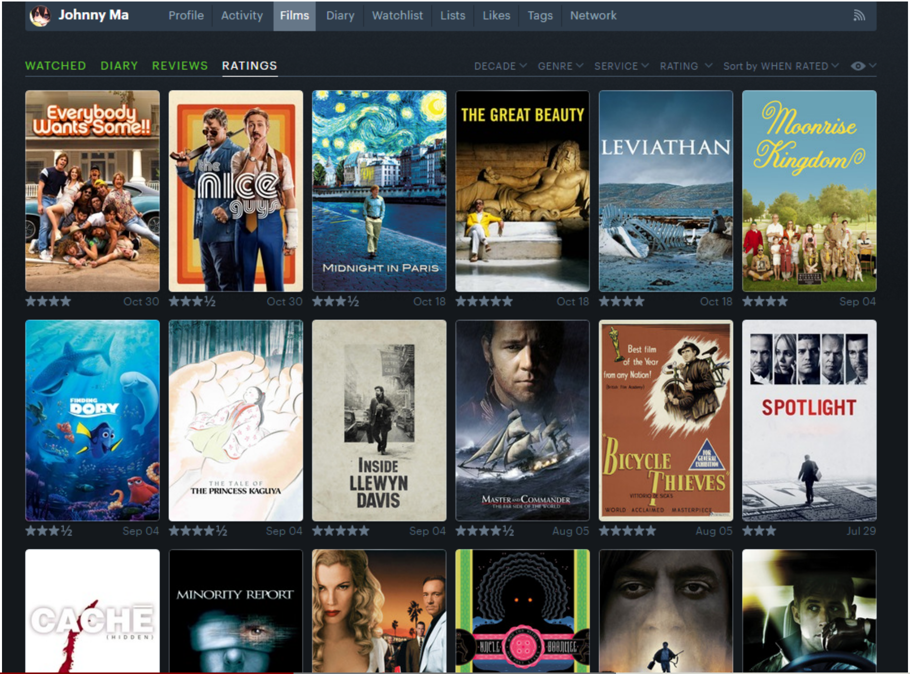

Johnny Ma
Profile
Activity
Films
Diary
Watchlist
Lists
Likes
Tags
Network
WATCHED
DIARY
REVIEWS
RATINGS
DECADE
GENRE
SERVICE
Sort by RELEASE DATE

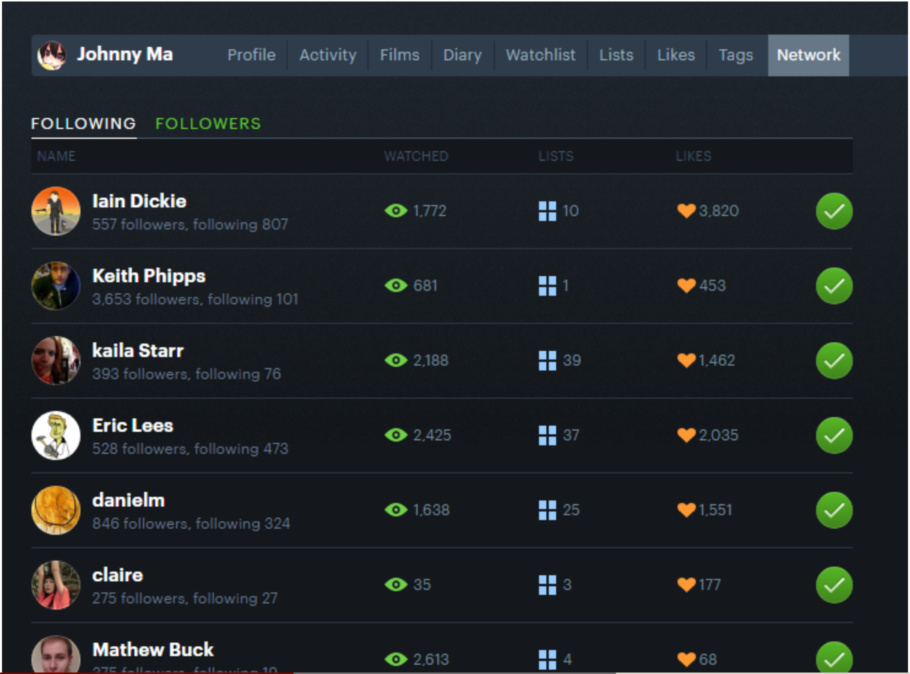

davidehrlich PATRON
New York City
indiewire.com
davidehrlich
FOLLOW
1,335 Films
133 This year
22 Lists
50 Following
16,738 Followers
Profile Activity Films Diary Watchlist Lists Likes Tags Network Stats
FAVORITE FILMS
CAROL
CLOSE-UP
A FILM BY ABBAS KIAROSTAMI
mishima
Lost In Translation
RECENT ACTIVITY
ALL
DAWSON CITY
FROZEN TIME
MUMMY
MEGAN
LEAVEY
BIO
senior film critic for Indiewire.
total moron.
STATS FOR 2017
Need upgrades?
Now with all-time stats!
Get PRO !
DIARY
1,123

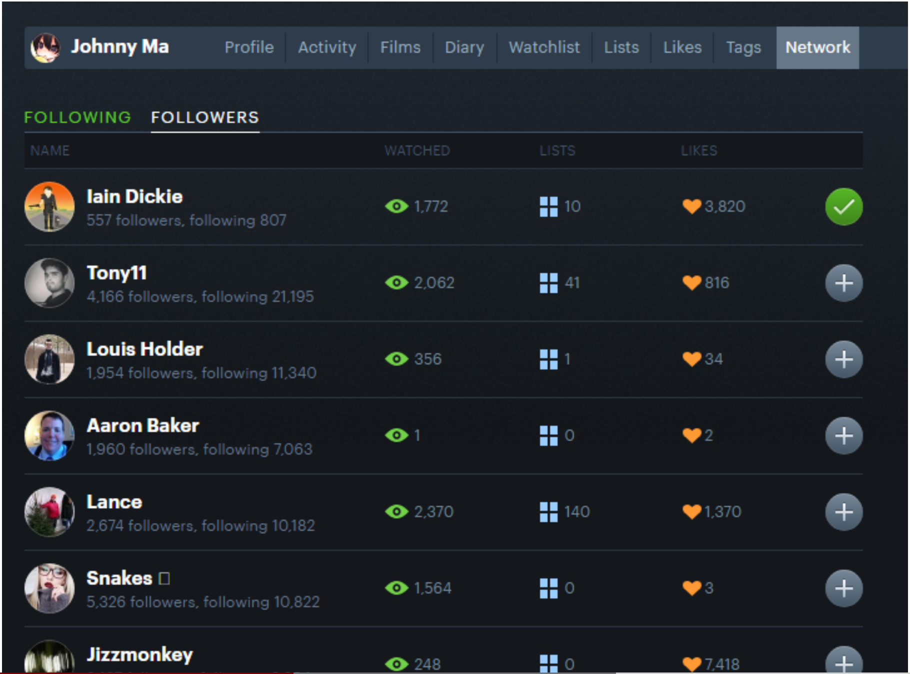

Adam Cloutier
Phoenix, AZ
vimeo.com/casualpunk
FOLLOWING
4,654 Films
155 This year
4 Lists
126 Following
496 Followers
Profile
Activity
Films
Diary
Watchlist
Lists
Likes
Tags
Network
FAVORITE FILMS
BNSF
BIO
vimeo.com/casualpunk
Canon/Favorites/whathaveyou
Need upgrades?
Now with all-time stats!
Get PRO !
RECENT ACTIVITY
ALL
TWIN PEAKS FIRE WALK WITH ME
HINTERLANDS
WATCHLIST
555
THE PIT AND THE PENDULUM


Jared S. PRO
WI jaredsletterbox
FOLLOW
2,018 Films
175 This year
50 Lists
136 Following
3,232 Followers
Profile Activity Films Diary Watchlist Lists Likes Tags Network Stats
FAVORITE FILMS
LORD OF THE RINGS
THE TREE OF LIFE
A L I E N
STAR WARS
RECENT ACTIVITY
ALL
WAR MACHINE
CAPTAIN AMERICA THE WINTER SOLDIER
CONTACT
WONDER WOMAN
RECENT REVIEWS
MORE
BIO
Avidly embracing nerd-culture and worshiping the Holy Trinity (Star Wars, Lord of the Rings, Star Trek) since 1997. I mostly do marathons. Here are my lists.
STATS FOR 2017
Need upgrades?
Now with all-time stats!
Get PRO !
WATCHLIST
299


Johnny Ma
Profile Activity Films Diary Watchlist Lists Likes Tags Network
WATCHED DIARY REVIEWS RATINGS
RATING YEAR DECADE GENRE SERVICE Sort by WATCHED DATE
MONTH DAY FILM RELEASED RATING LIKE REWATCH REVIEW EDIT
JUN 2017 01 Breathless 1960
MAY 2017 27 Your Name. 2016
20 Y Tu Mamá También 2001
APR 2017 03 Elle 2016
FEB 2017 14 Hunt for the Wilderpeople 2016
10 Hell or High Water 2016
10 La La Land 2016
10 Weiner 2016
10 The Handmaiden 2016


Johnny Ma
Profile
Activity
Films
Diary
Watchlist
Lists
Likes
Tags
Network
WATCHED
DIARY
REVIEWS
RATINGS
DECADE
GENRE
SERVICE
RATING
Sort by WHEN RATED
Everybody Wants Some!!
the nice guys
MIDNIGHT IN PARIS
THE GREAT BEAUTY
LEVIATHAN
Moonrise Kingdom
★★★★ Oct 30
★★★½ Oct 30
★★★½ Oct 18
★★★★★ Oct 18
★★★★ Oct 18
★★★★ Sep 04
FINDING DORY
THE TALE OF THE PRINCESS KAGUYA
INSIDE LLEWYN DAVIS
MASTER AND COMMANDER
THE FAR SIDE OF THE WORLD
Best film of the Year from any Nation!
BICYCLE THIEVES
WORLD ACCLAIMED MASTERPIECE
SPOTLIGHT
★★★½ Sep 04
★★★★½ Sep 04
★★★★★ Sep 04
★★★★½ Aug 05
★★★★★ Aug 05
★★★ Jul 29
CACHÉ
(HIDDEN)
MINORITY REPORT

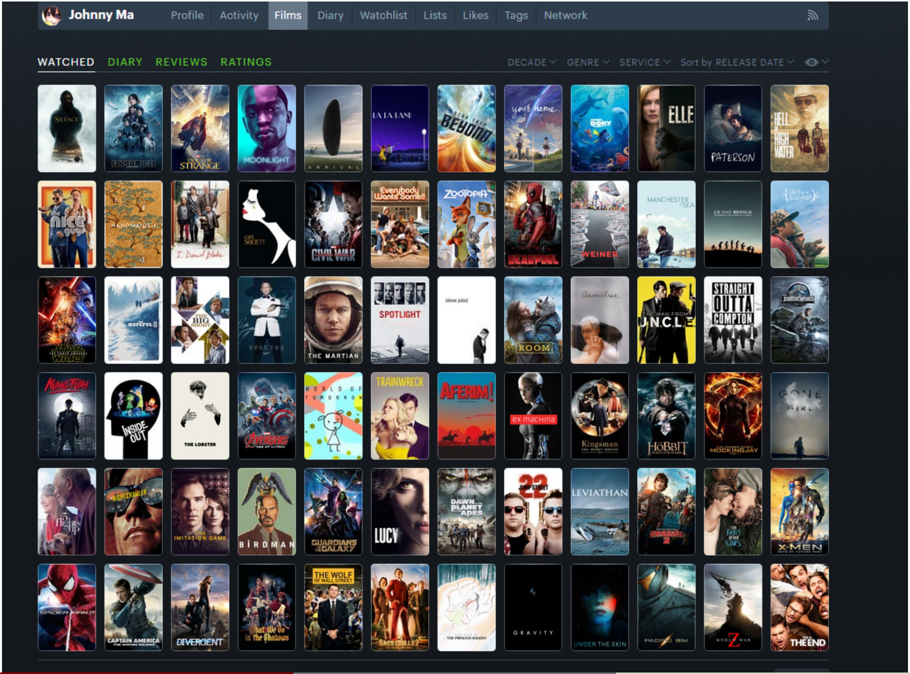

Johnny Ma
Profile
Activity
Films
Diary
Watchlist
Lists
Likes
Tags
Network
FOLLOWING
FOLLOWERS
NAME
WATCHED
LISTS
LIKES
Iain Dickie
557 followers, following 807
1,772
10
3,820
Keith Phipps
3,653 followers, following 101
681
1
453
kaila Starr
393 followers, following 76
2,188
39
1,462
Eric Lees
528 followers, following 473
2,425
37
2,035
danielm
846 followers, following 324
1,638
25
1,551
claire
275 followers, following 27
35
3
177
Mathew Buck
2,613
4
68



Johnny Ma | Profile | Activity | Films | Diary | Watchlist | Lists | Likes | Tags | Network

FOLLOWING FOLLOWERS

NAME	WATCHED	LISTS	LIKES
Iain Dickie 557 followers, following 807	1,772	10	3,820
Tony11 4,166 followers, following 21,195	2,062	41	816
Louis Holder 1,954 followers, following 11,340	356	1	34
Aaron Baker 1,960 followers, following 7,063	1	0	2
Lance 2,674 followers, following 10,182	2,370	140	1,370
Snakes □ 5,326 followers, following 10,822	1,564	0	3
Jizzmonkey	248	0	7,418


Johnny Ma
Profile
Activity
Films
Diary
Watchlist
Lists
Likes
Tags
Network
WATCHED
DIARY
REVIEWS
RATINGS
DECADE
GENRE
SERVICE
Sort by RELEASE DATE


the social network
The Social Network 2010 Directed by David Fincher
YOU DON'T GET TO 500 MILLION FRIENDS WITHOUT MAKING A FEW ENEMIES.
On a fall night in 2003, Harvard undergrad and computer programming genius Mark Zuckerberg sits down at his computer and heatedly begins working on a new idea. In a fury of blogging and programming, what begins in his dorm room as a small site among friends soon becomes a global social network and a revolution in communication. A mere six years and 500 million friends later, Mark Zuckerberg is the youngest billionaire in history... but for this entrepreneur, ...more
Watched
Like
Watchlist
Rated
Review or log
Add to a list
Share
118k 18k 31k
WATCH
Play trailer
Amazon US RENT BUY
iTunes US RENT BUY
MORE SERVICES
GoWatchIt
CAST CREW DETAILS GENRE
Jesse Eisenberg
Andrew Garfield
Justin Timberlake
Armie Hammer
Max Minghella
Rooney Mara
Brenda Song
Rashida Jones
John Getz
David Selby
Denise Grayson
Douglas Urbanski
Joseph Mazzello
Wallace Langham
Patrick Mapel
Dakota Johnson
Malese Jow
Bryan Barter
Barry Livingston
Shelby Young
Abhi Sinha
Mariah Bonner
Mark Saul
Cedric Sanders
Inger Tudor
Emma Fitzpatrick
John Hayden
James Shanklin
Oliver Muirhead
Show All...
120 mins More details at IMDB TMDB
RATINGS
3.9
WATCHED BY
118k people / 22 friends / 1.8k fans

POPULAR REVIEWS MORE

Review by **sree** ★★★★★ 14

me: i should get some sleep i have a lot to do in the morning
my brain: hey remember when eduardo's shares were diluted down to 0.03%?

Like this review? 835 likes

Review by **RagingTaxiDriver** ★★★★ 15

When you're an asshole, you lose your girlfriend.
When you lose your girlfriend, you get drunk at Harvard.
When you get drunk at Harvard, you rant on the Internet.
When you rant on the Internet, you make FaceMash.com.
When you make FaceMash.com, you crash the system.
When the system crashes, twins that row crew ask for your help (you hate crew).
When you hate crew, you steal their idea.
When you steal their idea, it becomes big.
When it becomes big, you become more of an asshole.
When you become more of an asshole, you lose your friends.
When you lose your friends, you get sued.
When you get sued, you lose millions of dollars.
When you lose millions of dollars, you try friending that ex-girlfriend on Facebook.

...don't be an asshole.

64k 12k 16k

WATCH

Play trailer

Amazon US RENT

Netflix US PLAY

MORE SERVICES GoWatchIt

REVIEWS FROM FRIENDS MORE

Review by **Matt Singer** ★★★½ 2

The evolution of Adam McKay. From Will Ferrell teabagging drum sets, to entitled douchebag financial bros metaphorically teabagging America.

Full review: **screencrush.com/the-big-short-review/**

Like this review? 52 likes

Review by **Matt Singer** ★★★½ 8

The stuff that didn't work about this the first time really didn't work for me the second time, mostly Steve Carell's performance and the script's fumbling attempts to give him a nagging conscience and a tragic backstory. But the stuff that did work continued to work, and man amongst everything else that's good about this movie there are some lines in here as great and quotable as anything McKay's ever done (the one about the line between illegal and stupidity is uh maze ing). Plus it dawned on me this time that Michael Burry's drum set plays the same rule as Dale Doback's drum set in STEP BROTHERS: As the ultimate symbol of defiant, puffed-up masculinity.

Like this review? 46 likes

Review by **Mr. DuLac** ★★★★ 1

From the director of FIVE ***Will Ferrell*** star vehicles comes a film based on the 2008 financial collapse primarily caused by the housing market. It's one of the best, funniest and most entertaining films of the year and practically a shoe in for a Best Picture Oscar nomination. None of that is sarcasm, and yet all those elements really don't belong together.

Like this review? 26 likes

Methodology 1

Though nice for web scraping, many issues with specific users and films.

Dropped those with no friends, dropped those with no diary, dropped those with no ratings, etc.

Chose 5000 users (from popular users list) and 720 films (from popular films list). Easily extend-able, but this is what my computer can handle. As long as more users than films and not putting more than just dummies into the regression, should avoid full-identification and collinearity issues.

Methodology 2

Created the "surprise" by calculating average rating for user i and setting 1 if film j is above, and 0 if below.

First 500 friends are taken. Most don't have 500 (32 friend average, many large outliers from 0 to 4,000). Count only friends who have seen and rated film j in the "surprise" share ratio.

Chose a period from 2017-01-01 to 2017-06-01, for recency and convenience. Easily changeable, not much time to test multiple.

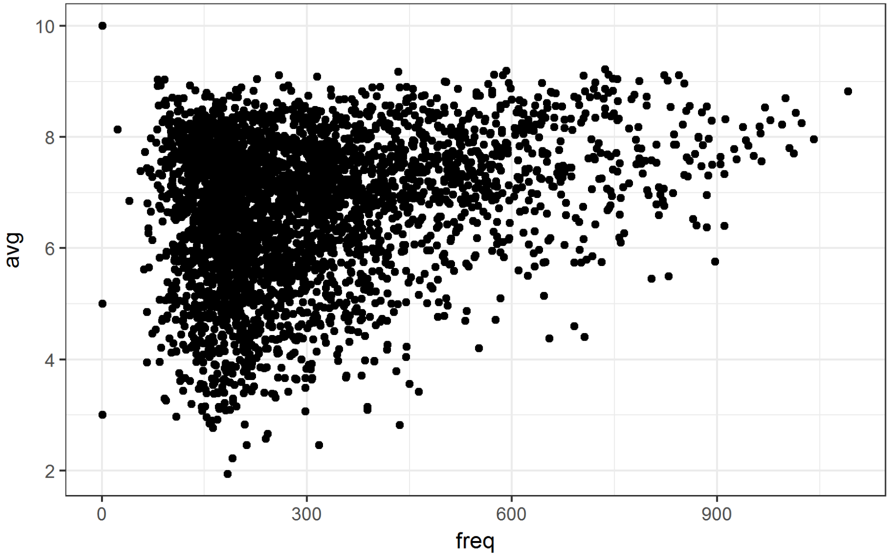
Data Conclusion

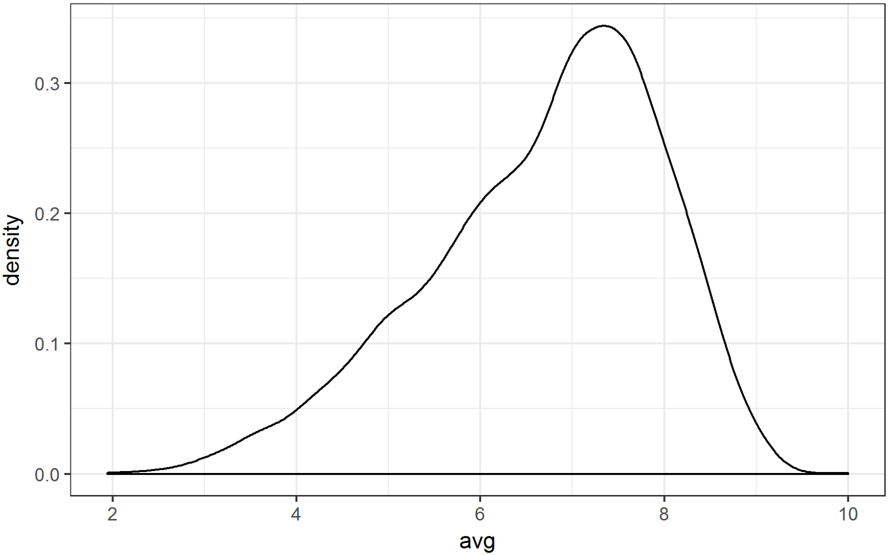
The best possible data to do microfounded studies into film watching and rating behavior. We will be focusing mostly on watching.

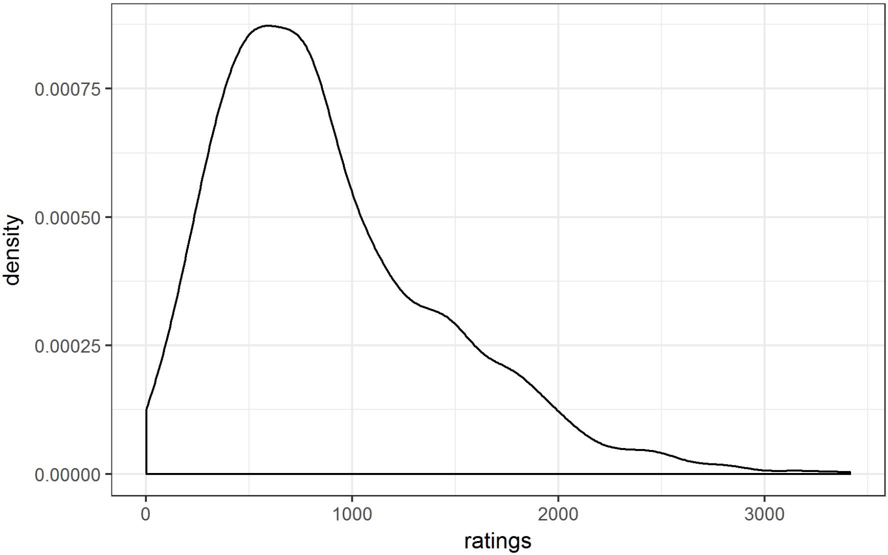
Still potential variation unexplained by regression, but in aggregate should work out well.

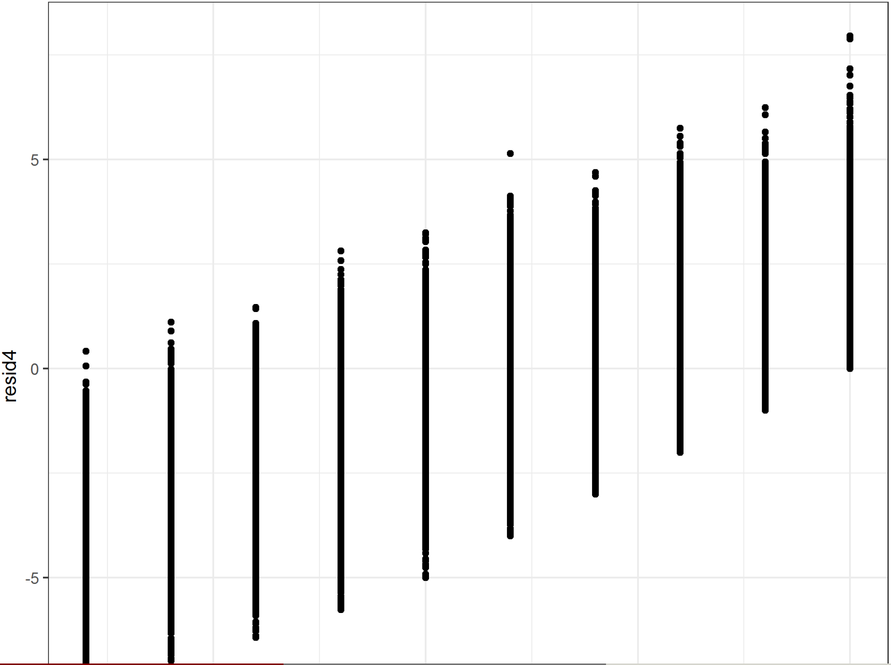
To be honest, I don't know how others use this site. Aside from summary statistics, it is difficult to tell.

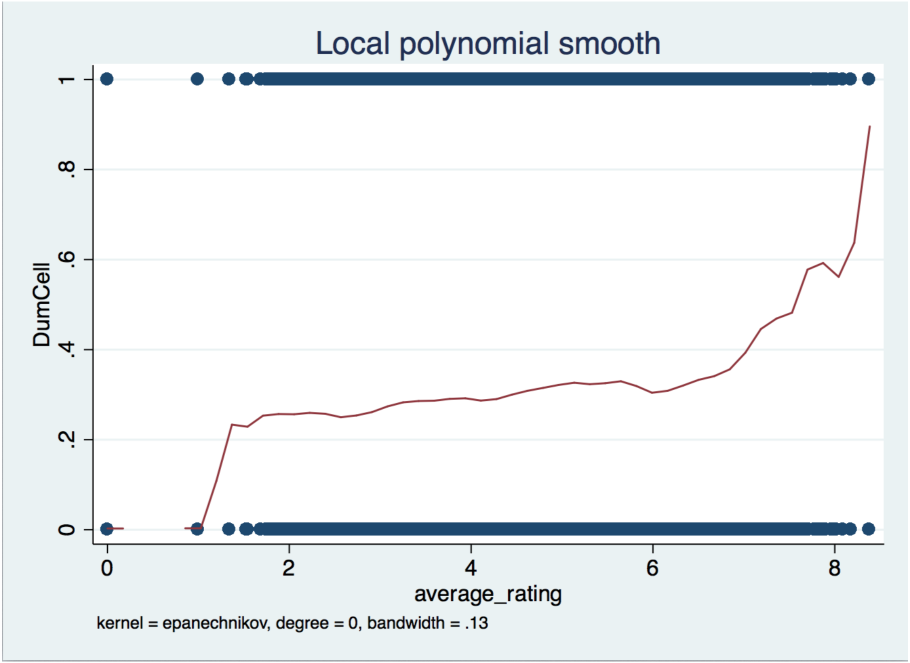
Wrote many flexible functions to scrape and organize data. Set of functions a great tool for specifying parameters to gather and look at data. Collected a lot of user and film data that was not used due to time constraints.


10
8
6
4
2
avg
0
300
600
900
freq


density
0.0
0.1
0.2
0.3
2
4
6
8
10
avg


density
0.00075
0.00050
0.00025
0.00000
0
1000
2000
3000
ratings


resid4
5
0
-5


Local polynomial smooth
DumCell
0
.2
.4
.6
.8
1
0
2
4
6
8
average_rating
kernel = epanechnikov, degree = 0, bandwidth = .13

Regression Results

Table: Multiple Group Fixed Effects Model

	Dependent variable:			
	Watched in 2017 If Not Seen Already			
	(1)	(2)	(3)	(4)
shares	0.062***	0.052***	0.055***	0.047***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Constant	0.024***			
	(0.0003)			
Fixed effcts?	No	User	Film	Both
Observations	903,700	903,700	903,700	903,700
R^2	0.024	0.035	0.031	0.040
Adjusted R^2	0.024	0.035	0.031	0.040
Residual Std. Error	0.210 (df = 903698)	0.208 (df = 903696)	0.209 (df = 903696)	0.208 (df = 903694)

Note: *p<0.1; **p<0.05; ***p<0.01

Discussion

Found that there is some correlation between friends who are "surprised" by a movie and an individual's probability of watching the film (if they have not watched it before).

Can't say much about identification, but can say that there is a correlation: either networks form around certain "tastes" or people see their friends' ratings and take that into account.

Another explanation is that each network is experiencing some group-level shock: everyone goes and sees the new movies in theatres, biasing trend upwards.

Any thoughts?

Extensions

Simply endless. More data, more users, more films (including older, less popular).

Using the user-favorite data along with film characteristics to increase R-squared and look more into what makes films popular.

Estimation of some social multiplier. Need a better sense of how these networks form and the relative weights.

Exploit how friends show up first on reviews, then popular reviews. Variation on what you see can be used conditional on your friend groups. Which one matters more? Other questions of identification, need exogenous variation or IV.

Using more outside information like box-office returns, geographic releases, etc.

Acknowledgements

Thanks to Professor Manresa for helping on this project and for teaching this wonderful class. I thought a lot more about identification than I ever have.

Thanks to Sylvia Klosin for help in building the model, Paul Beckman for intuition about films, and William Jones for general moral support.

Thanks to Letterboxd.com for being nice and not banning my IP when I made 5000 requests in one second. Also, for building such a wonderful platform!.

Thanks to you, for listening and being interested!