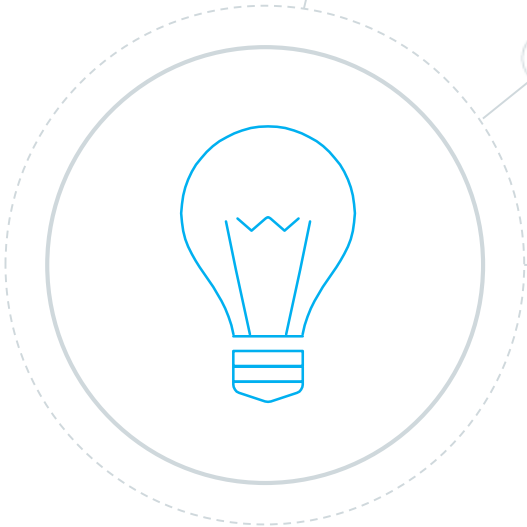
Project Luther

Too many movies?

Situation

- Luther Films plans to release a film in July 2016 but they found out that there will be a number of similar (same-genre) movies that will be released within +/- 30 days ("density")
- Luther funded the production with high-yield debt and is worried that the competition might affect its ability to service the debt. On the other hand, interest will accrue longer if the release is delayed
- The managers at Luther Films reached out to Metis to help them on this issue and that Metis come back with a recommendation the next day

The idea



Competition decreases revenue, or does it? The average American watches a movie every 2 months

L.E.K Box Office Research

Approach

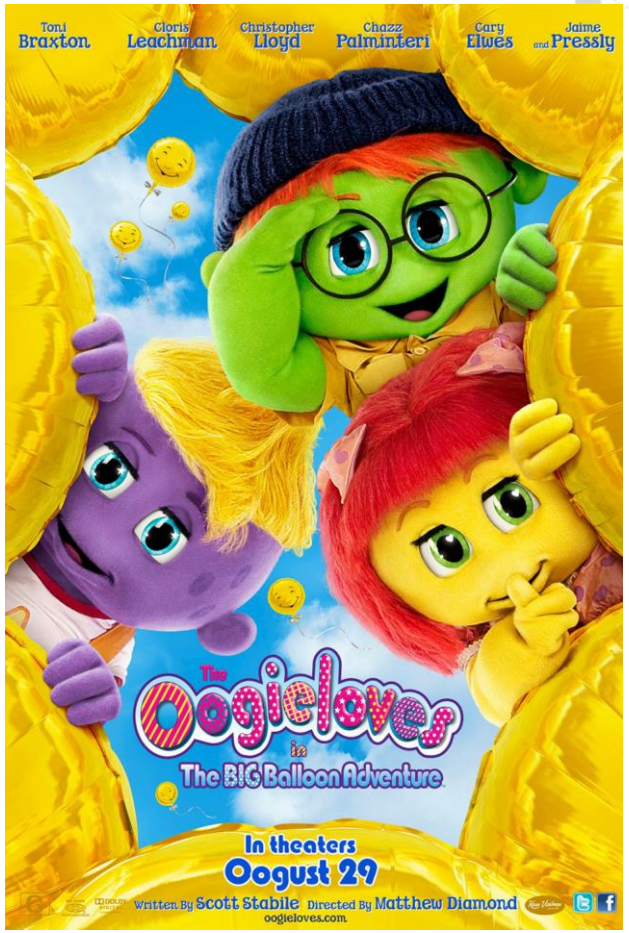
- Given the time and budget constraints…
- Data collection through web scraping
- Modeling through linear regression

Data Munging

- Data: Movies from 2009 to 2015
- Data source: BoxOfficeMojo
- Webpages scraped: 4,400
- Missing data treatment: Complete cases only
- Filter out: removed blockbusters, foreign movies, any movie with box office gross and budget below US$1 million
- Centered “# of theaters” and “# of days in release” variables to address severe multicollinearity
- Collapsed “genre” and “distributor” categorical variables
- Final # of observations: 730

Beware of the Outliers

- Distributed by Kenn (Viselman)
- Domestic gross of US$1 million
- Budget of US$20 million
- While not statistically significant, OLS shows that getting Kenn as the distributor increases domestic box office gross by US$48 million relative to “Other Distributors”



Source: BoxOfficeMojo

Regression Model

- Evaluated OLS and Elastic Net Regression
- Used 100% of data since focus in on coefficient
- Model equation:

adj. dom. gross = 1 + adj. budget + runtime + # days in release + max # of theaters (C) + (max # of theaters (C))^2 + density + genre + distributor + month

	R^2	Adj. R^2	Density Coefficient p-value
Ordinary Least Squares	**73%**	**73%**	**6%**
Elastic Net Regression	**72%**	**72%**	**21%**

US$ 2,200,000

less revenue for every same-genre movie that will be released within 30 days Luther Film's movie release date ($p < 0.07$, HC3)

Luther Films should consider releasing the movie in June, a month earlier, to catch the spring/summer rush and potentially increase revenue by **US$15 million** *($p < 0.09$, HC3)*

Other Ideas

- Density measured by # of high budget movies and (just the) # of movies released within +/- 30 days
- Scraped OMDB data for movie ratings from Rotten Tomatoes, IMDB, Metacritic
- Scraped RGB pixel data of movie posters

Thanks!

Any questions?