



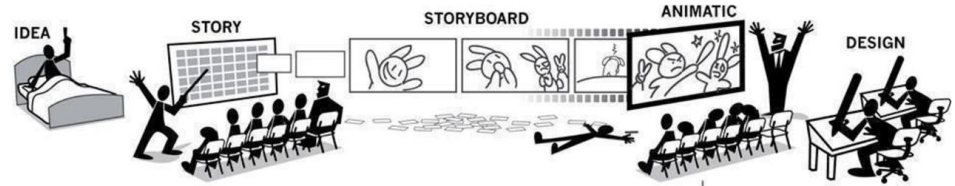
Predicting Revenue for Animation Films

Tiffany Li

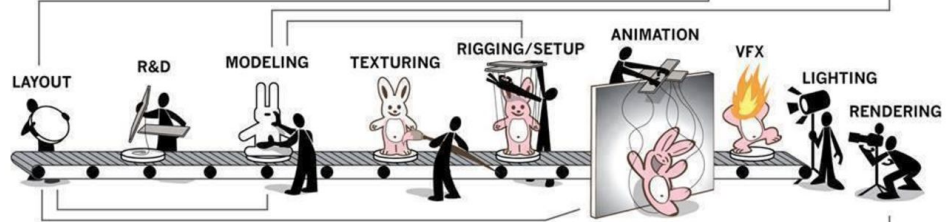
Motivations

- The process for producing animation films is long, complex and expensive
- Joint efforts of up to 500 talents
- What are the keys to success?

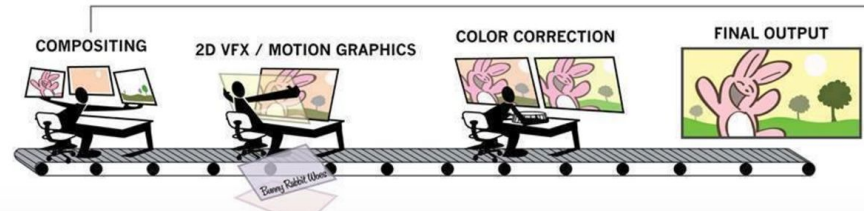
PRE-PRODUCTION



PRODUCTION



POST-PRODUCTION



Web Scrapping

Data collected from imdb.com

Genres: Animation | Adventure | Comedy | Family

Motion Picture Rating (MPAA)
Rated PG for mild thematic elements | See all certifications »
Parents Guide: View content advisory »

Details

Official Sites: Official Facebook | Official site | See more »
Country: USA
Language: English | Indonesian
Release Date: 17 June 2016 (USA) See more »
Also Known As: Finding Nemo 2 See more »
Filming Locations: Pixar Animation Studios - 1200 Park Avenue, Emeryville, California See more »

Box Office

Budget: \$200,000,000 (estimated)
Opening Weekend: \$135,060,273 (USA) (17 June 2016)
Gross: \$486,292,984 (USA) (2 December 2016)
See more »

Company Credits

Production Co: Pixar Animation Studios, Walt Disney Pictures See more »
Show detailed company contact information on IMDbPro »

Technical Specs

Runtime: 97 min
Sound Mix: Dolby Digital | Datasat | Dolby Atmos | SDDS | DTS
Color: Color
Aspect Ratio: 1.85 : 1
See full technical specs »

```
Elements Console Sources >> 4 X
<!DOCTYPE html>
<html xmlns:og="http://ogp.me/ns#" xmlns:fb="http://www.facebook.com/2008/fbml" class="scriptsOn gr_imdb.com">
  <head>...</head>
  <body id="styleguide-v2" class="fixed data-gr-c-s-loaded="true">
    <script>
      if (typeof uet == 'function') {
        uet("bb");
      }
    </script>
    <script>...</script>
    <div id="wrapper">
      <div id="imdb-ptp-wrap">
        <span class="imdb-ptp-wrap imdb-ptp-wrap-active"><span>
          <span class="imdb-ptp-wrap transition" style="background-image: url('http://g-ec2.images-amazon.com/images/G/01/IMDb/ads/2016/PTP/24893-FindingDory/ptp-wrap-videos-4.jpg')">
            </span>
          </div>
        <div id="root" class="redesign">...</div>
        <script type="text/javascript" src="http://ia.media-imdb.com/images/G/01/imdb/is/collections/common-485701321_CB522736548.js"></script>
        <script type="text/javascript" src="http://ia.media-imdb.com/images/G/01/imdb/is/collections/title-2918568030_CB534889535.js"></script>
        <div id="photo-container">...</div>
        <div id="video-container">...</div>
        <script type="text/javascript" id="login">...</script>
      </div>
    </div>
  </body>
</html>
```



Top-US-Grossing Animation Feature Films

1 to 50 of 5,035 titles | Next »

View Mode: Compact | Detailed

Sort by: Popularity | Alphabetical | IMDb Rating | Number of Votes | **US Box Office ▼** | Runtime | Year | Release Date



1. Finding Dory (2016)

PG | 97 min | Animation, Adventure, Comedy

★ 7.5 this Rate 77 Metascore

The friendly but forgetful blue tang fish begins a search for her long-lost parents, and everyone learns a few things about the real meaning of family along the way.

Directors: Andrew Stanton, Angus MacLane | **Stars:** Ellen DeGeneres, Albert Brooks, Ed O'Neill, Kaitlin Olson

Votes: 129,913 | Gross: \$486.29M



2. Shrek 2 (2004)

PG | 93 min | Animation, Adventure, Comedy

★ 7.2 this Rate 75 Metascore

Princess Fiona's parents invite her and Shrek to dinner to celebrate her marriage. If only they knew the newlyweds were both ogres.

Directors: Andrew Adamson, Kelly Asbury, Conrad Vernon | **Stars:** Mike Myers, Eddie Murphy, Cameron Diaz, Julie Andrews

Votes: 327,140 | Gross: \$436.47M



3. The Lion King (1994)

G | 88 min | Animation, Adventure, Drama

★ 8.5 this Rate 83 Metascore

Lion cub and future king Simba searches for his identity. His eagerness to please others and penchant for testing his boundaries sometimes gets him into trouble.

Directors: Roger Allers, Rob Minkoff | **Stars:** Matthew Broderick, Jeremy Irons, James Earl Jones, Whoopi Goldberg

Votes: 679,392 | Gross: \$422.78M

Data Processing

Using **beautiful soup**, scraped features for top 400 animation films, including movie title, budget, country, MPAA rating, opening weekend gross, production co, IMDB rating score, release date, runtime and awards.

Awards	Budget	Certificate	Color	Country	Filming Locations	Gross	Language	...	Movie Title	Official Sites	Opening Weekend
[' Nominated for 1 BAFTA Film Award.', ' A...	\$200,000,000	NaN	Color	USA	Pixar Animation Studios - 1200 Park Avenue, Em...	\$486,292,984	English	...	Finding Dory (2016) - IMDb	Official Facebook	\$135,060,273
[' Nominated for 2 Oscars.', ' Another16 w...	\$150,000,000	NaN	Color	USA	NaN	\$436,471,036	English	...	Shrek 2 (2004) - IMDb	NaN	\$108,037,878

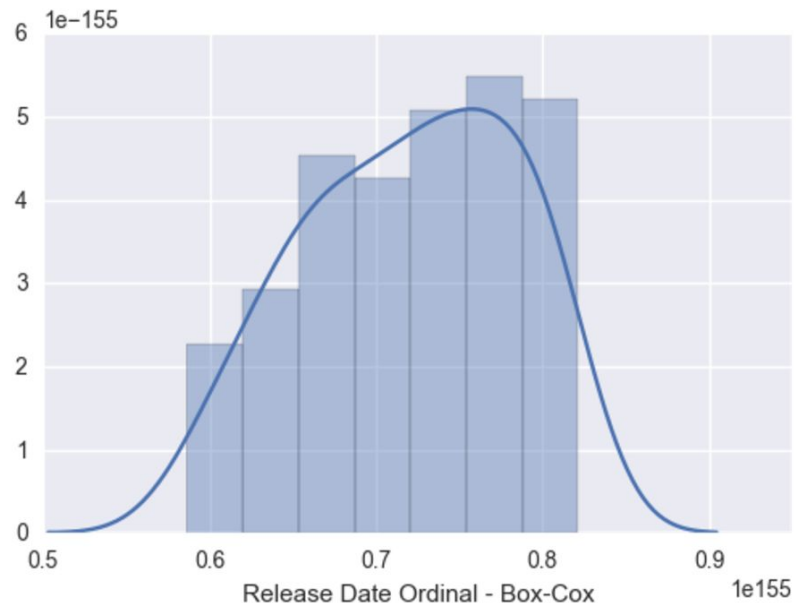
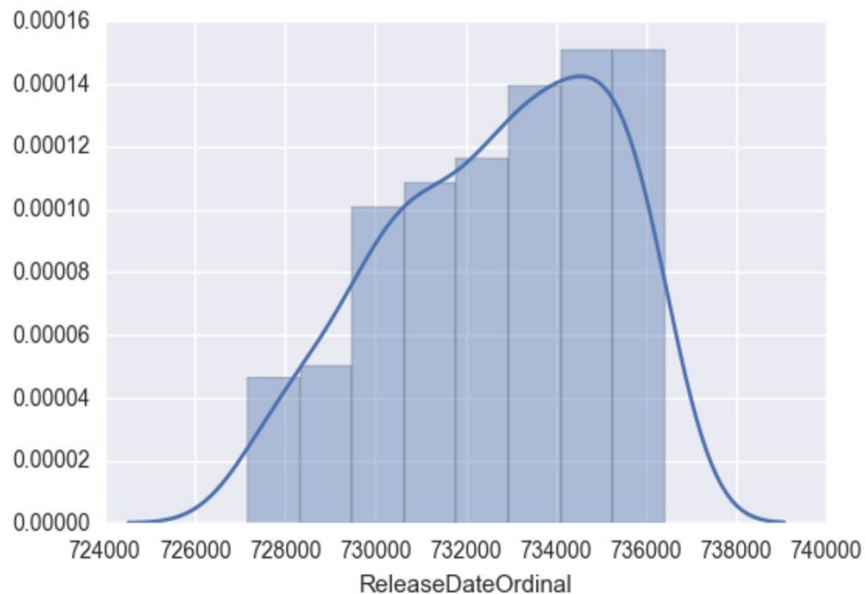
Feature Engineering

- Dropped data outside of US or earlier than 1990
- Categorized movies into 'won awards', 'nominated' or 'not'
- Categorized production companies into large, medium and small based on total gross revenue
- Extracted year, month, day, day of the week and date ordinal
- Filled in missing value with median



223 data points as input

Box-Cox Transformation



Finding the Best Model

- Used simply linear regression to compare effectiveness by adding/dropping variables
- Found the best predictor variables to be **opening weekend**, **rating score**, **release date ordinal**, **awards results** and **production company size**.

OLS Regression Results

Dep. Variable:	GrossM	R-squared:	0.831
Model:	OLS	Adj. R-squared:	0.826
Method:	Least Squares	F-statistic:	176.9
Date:	Thu, 02 Feb 2017	Prob (F-statistic):	1.81e-80
Time:	22:10:56	Log-Likelihood:	-1150.0
No. Observations:	223	AIC:	2314.
Df Residuals:	216	BIC:	2338.
Df Model:	6		
Covariance Type:	nonrobust		

Regularization

- Compared cross validated Lasso, Rigid and Elastic Net models
- **Cross validated Elastic Net model** produced the best mean squared error
 - Alpha = 0.17
 - R squared = 0.878

Predictive Model

Gross Revenues Equals to

$$96 + 62 * \text{Opening Weekend} + 14 * \text{Rating Score} + 0.0 * \text{Release Date Ordinal} \\ + 4 * \text{Nominated} + 42 * \text{Won} + 21 * \text{Large} - 23 * \text{Small}$$

Example: Inside out

Model predicts: 340M



Predictive Model

Gross Revenues Equals to

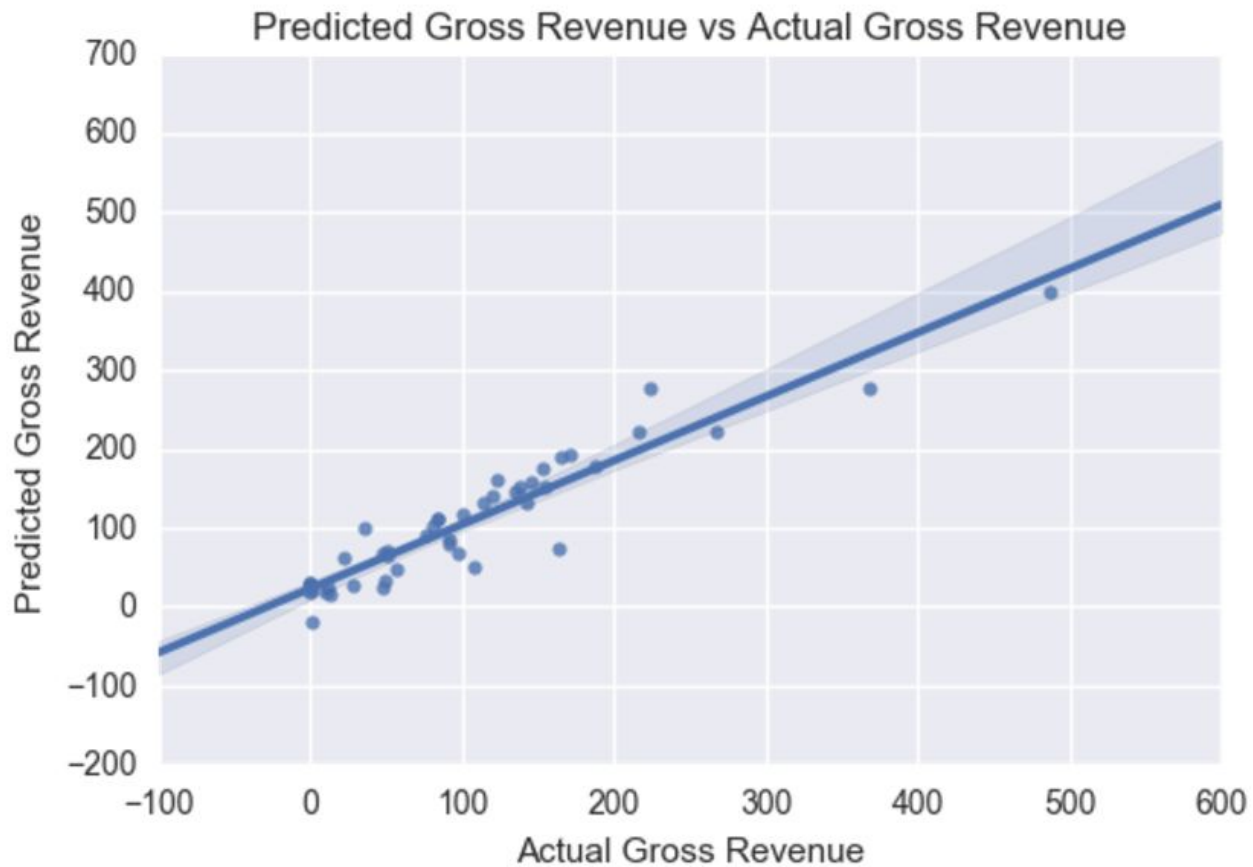
$96 + 62 * \text{Opening Weekend} + 14 * \text{Rating Score} + 0.0 * \text{Release Date Ordinal} + 4 * \text{Nominated} + 42 * \text{Won} + 21 * \text{Large} - 23.4 * \text{Small}$

Example: Inside out

Model predicts: 340M

Actual: 356M





Implications

What's important?

- Opening Weekend
 - First impression matters! Put out more commercials
- IMDB Rating Score
 - Get that 'critically acclaimed' stamp
- Awards
 - Oscar, here I come...
- Production Company Size
 - Big studios still got the edge



Implications

What's not worth it:

- Budget
 - Maybe don't spend 6 months making a snowflake?
- Release Date Month / Quarter
 - Christmas doesn't guarantee the presents
- MPAA Rating
 - G or PG? Does not matter



In an effort to perfect Elsa's icy magic in Frozen, filmmakers called on Dr. Thomas Painter to learn about snowflakes from a molecular level.

Further Investigations

- More feature engineering
 - Break down the awards into different levels
- Incorporate data from other sources
 - Fill in missing values with actual values
 - Other rating metrics
 - Impact of 3D movies
 - Does the popularity of the voice over cast matter?



*A minion
thanks.*