

# PREDICTING THE NUMBER OF MOVIE TORRENTS TO APPEAR ONLINE

BRYANT BIGGS

15<sup>TH</sup> JUNE 2016

**Best Torrent  
Sites For Movies**



ExtraTorrent

RARBG.com

1337X  
TORRENTS

ISOHUNT

BitSnoop

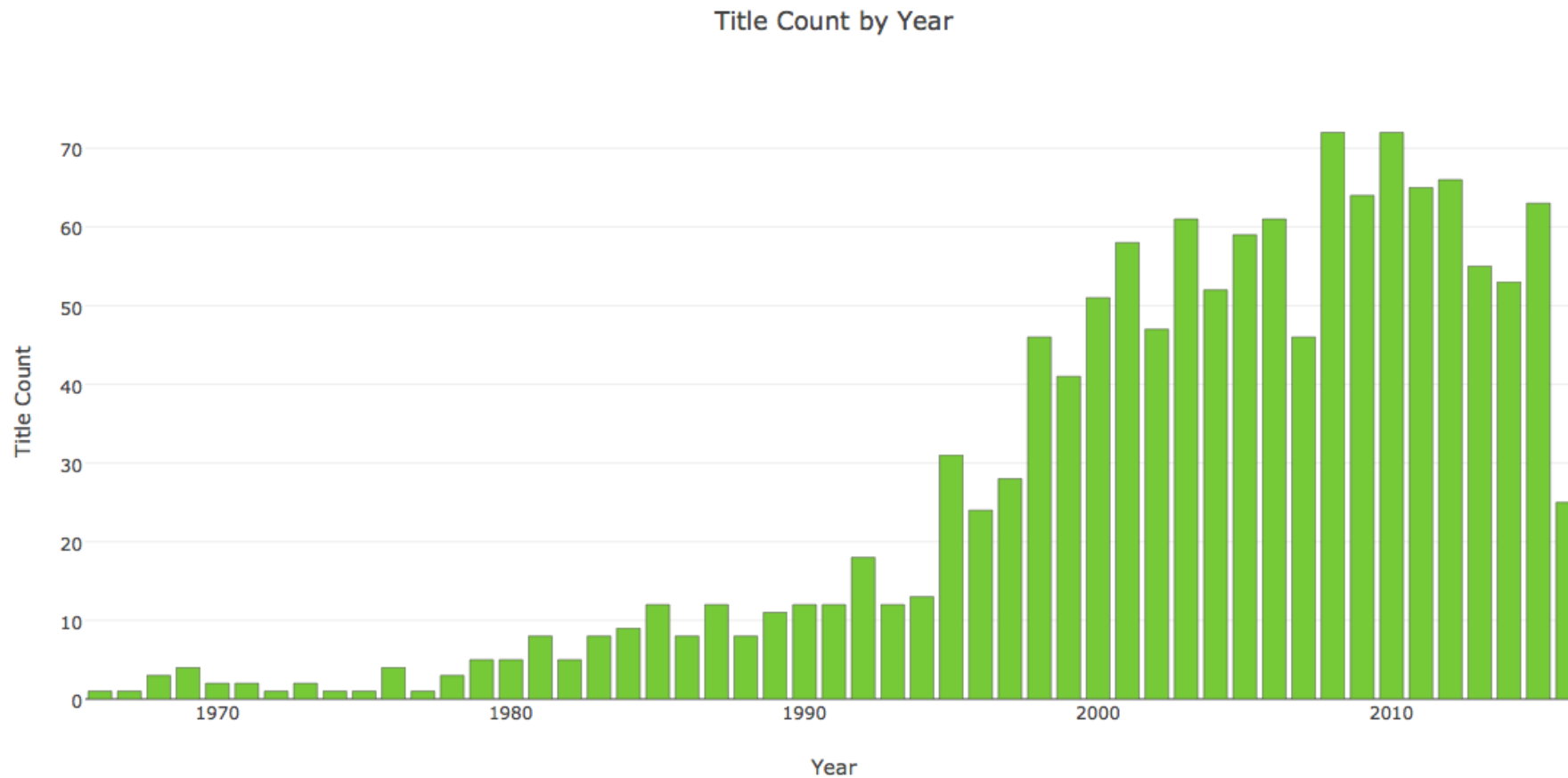
# DATA

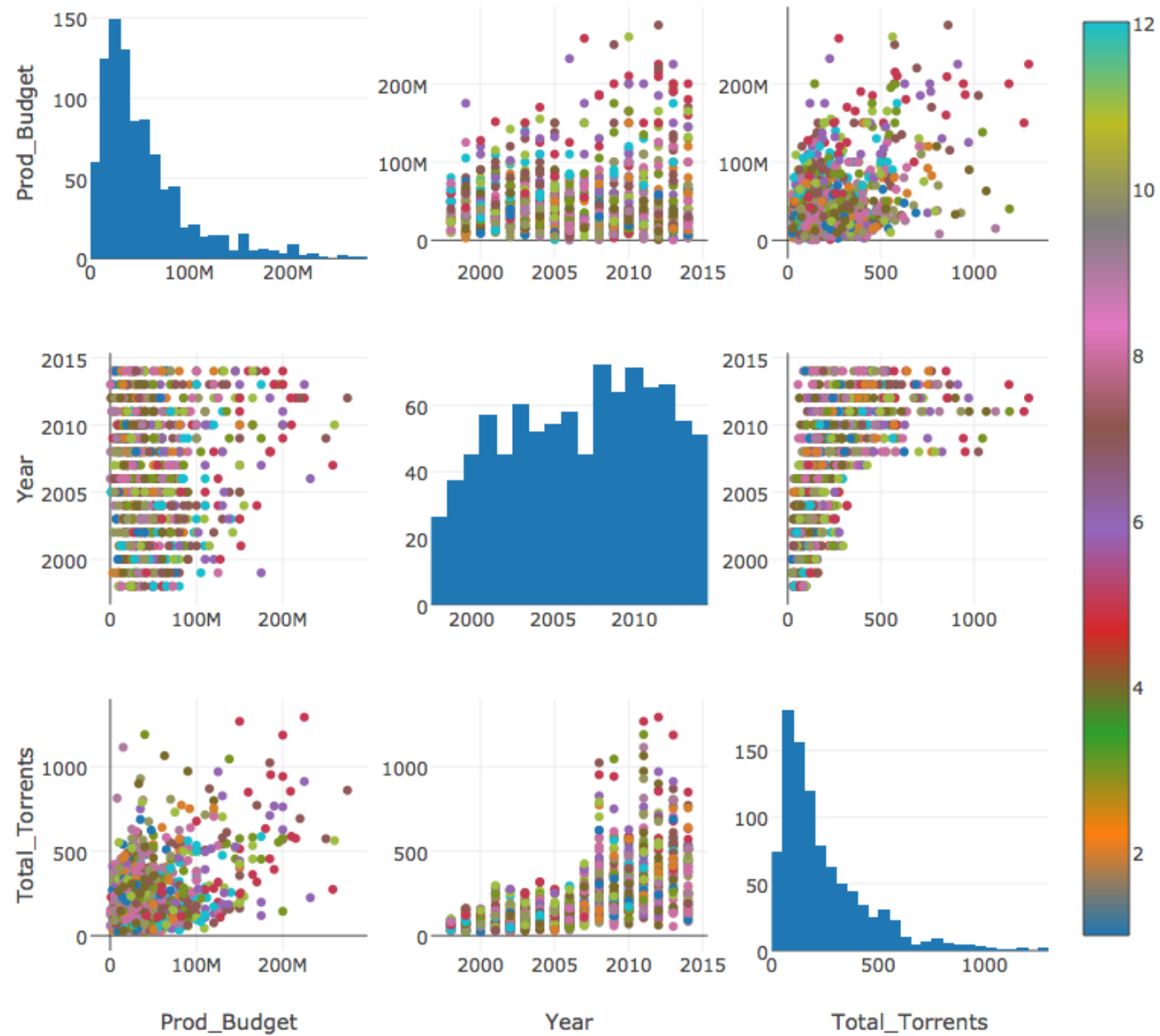
- `www.the-numbers.com` => Title & Production Budget
- `www.omdbapi.com` => Genre, Release Date, Runtime, Rating, Actors, Runtime, etc.
- `torrentz.eu` => Torrent Count
- `kat.cr` => Torrent Count
- `thepiratebay.org` => Torrent Count

	Title	Budget	Released	Year	Month	Rated	Runtime	Genre	Director	Actors	Torrents
0	The Siege	70000000.0	1998-11-06	1998	11	R	116	Action, Thriller	Edward Zwick	Denzel Washington, Annette Bening, Bruce Willi...	113

# CUTOFFS

- **Years:** 1998 – 2015
- **Genres:** – Action, Adventure, Comedy, Drama





STANDARD



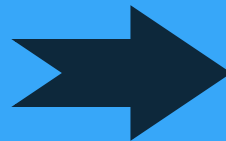
LOG TRANSFORM

# RESULTS

<b>R-squared:</b>	0.638
<b>Adj. R-squared:</b>	0.635
<b>F-statistic:</b>	181.3
<b>No. Observations:</b>	831
<b>Df Residuals:</b>	822

<b>Features</b>	<b>coef</b>
Genre[T.Adventure]	-0.2901
Genre[T.Comedy]	-0.5138
Genre[T.Crime]	-0.6252
Genre[T.Drama]	-0.3233
log_budget	0.2505
Year	0.1249

**Train Model Score = 0.64**  
**Train Mean Square Error = 0.26**



**Test Model Score = 0.69**  
**Test Mean Square Error = 0.25**

# FUTURE

- - Reduce data fallout
- - Explore other analysis models & regularization
- - Collect data on torrent quality
- - Time series analysis with in inflation
- - Better classification for genre (varies by website)

Thank you.