
Movie Analysis

Budget vs Domestic Gross for
Movies with a \$30M Budget

by Jason Keung

**Question: Given a genre and budget of 30 million dollars or less,
can I predict gross domestic profit using linear regression?**

Where did the data come from?

Search Site

Social

[Facebook](#)
[Twitter](#)

Features

[News](#)
[Release Sched.](#)
[Showtimes](#)
[at !\[\]\(758ebdf4629c903da74c2e079717ae32_img.jpg\)](#)

Box Office

[Daily](#)
[Weekend](#)
[Weekly](#)
[Monthly](#)
[Quarterly](#)
[Seasonal](#)
[Yearly](#)
[All Time](#)
[Chart Watch](#)
[International](#)

Indices

[Movies A-Z](#)
[Studios](#)

> MOVIES

ALPHABETICAL INDEX

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

[A-Ac](#) | [Ad-Ae](#) | [Af-Ag](#) | [Ah-AI](#) | [Am-Am](#) | [An-An](#) | [Ap-Ar](#) | [As-At](#) | [Au-Ay](#) | [Az-Az](#)

Title (click to view box office)		Studio	Total Gross / Theaters		Opening / Theaters		Open
The A-Team		Fox	\$77,222,099	3,544	\$25,669,455	3,535	6/11/2010
A.C.O.D.		TFA	\$175,705	42	\$19,001	3	10/4/2013
A.I. Artificial Intelligence		WB	\$78,616,689	3,242	\$29,352,630	3,242	6/29/2001
Aaja Nachle		Yash	\$484,108	66	\$257,500	66	11/30/2007
Aarakshan (Reservation)		Relbig.	\$651,096	91	\$344,661	91	8/12/2011
Aashayein (Wishes)		Relbig.	n/a	14	n/a	14	8/27/2010
Aatsinki: The Story of Arctic Cowboys		Argo.	\$6,601	2	\$2,842	1	1/24/2014
Abandon		Par.	\$10,727,683	2,347	\$5,064,077	2,341	10/18/2002
The Abandoned		ADF	\$1,331,137	1,000	\$782,000	1,000	2/23/2007
ABC Africa		NYer	\$9,075	1	\$559	1	5/3/2002
The ABC's of Death		Magn.	\$21,832	19	\$10,110	19	3/8/2013
ABCD		Eros	\$83,372	38	\$83,372	38	11/30/2001

Box Office Mojo - Individual Movie Page

Box Office Mojo

Adjuster: Actuals Go

Search Site
Search...

Social
[Facebook](#)
[Twitter](#)

Features
News
Release Sched.
Showtimes
at [IMDb](#)

Box Office
Daily
Weekend
Weekly
Monthly
Quarterly
Seasonal
Yearly
All Time
Chart Watch
International



Transformers: Age of Extinction

Domestic Total Gross: **\$245,439,076**

Distributor: **Paramount** Release Date: **June 27, 2014**

Genre: **Sci-Fi Action** Runtime: **2 hrs. 45 min.**

MPAA Rating: **PG-13** Production Budget: **\$210 million**

[Summary](#) [Daily](#) [Weekend](#) [Weekly](#) [Foreign](#) [Similar Movies](#) [Images](#)

Total Lifetime Grosses
Domestic: **\$245,439,076** 22.2%
+ Foreign: **\$858,600,000** 77.8%
= Worldwide: **\$1,104,039,076**

Domestic Summary
Opening Weekend: \$100,038,390 (#1 rank, 4,233 theaters, \$23,633 average)
% of Total Gross: 40.8% > View All 15 Weekends
Widest Release: 4,233 theaters

The Players
Director: Michael Bay
Writer: Ehren Kruger
Actors: Mark Wahlberg, Nicola Peltz, Jack Reynor, TJ Miller, Stanley Tucci, Kelsey Grammer
Producers: Michael Bay, Connelly, ...

Images

> View All 38 Images

Related Stories
9/10/14 Grading Mojo's Summer 2014 Forecast

Data Processing and Transformation Assumptions

Filter For Movies Under 30 Million Budget

15,000 Movies
(Any Budget)



1,500 Movies
(< \$30M Budget)

Removed Data Without Missing Values

Box Office Mojo

Search Site

Search...

Social

[Facebook](#)

[Twitter](#)

Features

News

Release Sched.

Showtimes

at [IMDb](#)

Box Office

Daily

Weekend

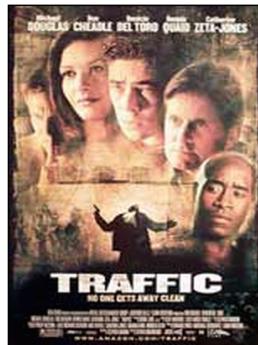
Weekly

Monthly

Quarterly

Seasonal

Yearly



[Summary](#)

[Daily](#)

[Weekend](#)

[Weekly](#)

[Foreign](#)

[Similar Movies](#)

Total Lifetime Grosses

Domestic: **\$124,115,725** 59.8%
+ Foreign: \$83,400,000 40.2%

= Worldwide: **\$207,515,725**

Domestic Summary

Release Date: December 27, 2000

The Players

Director: Steven Soderbergh
Actors: Michael Douglas
Don Cheadle
Benicio Del Toro
Dennis Quaid
Catherine Zeta-Jones

Related Stories

4/26/10 'It's Complicated' Debuts, 'Out of Africa,' 'Tombstone' Go Blu

Related Products

- Traffic
- Traffic

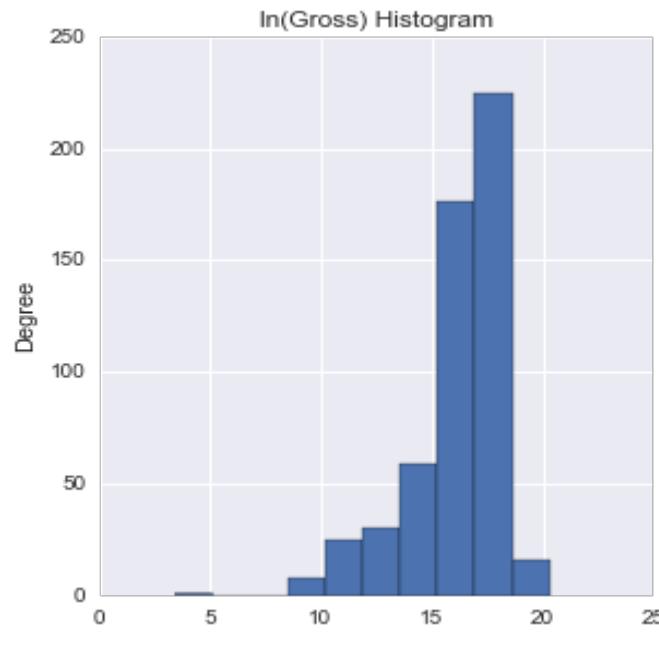
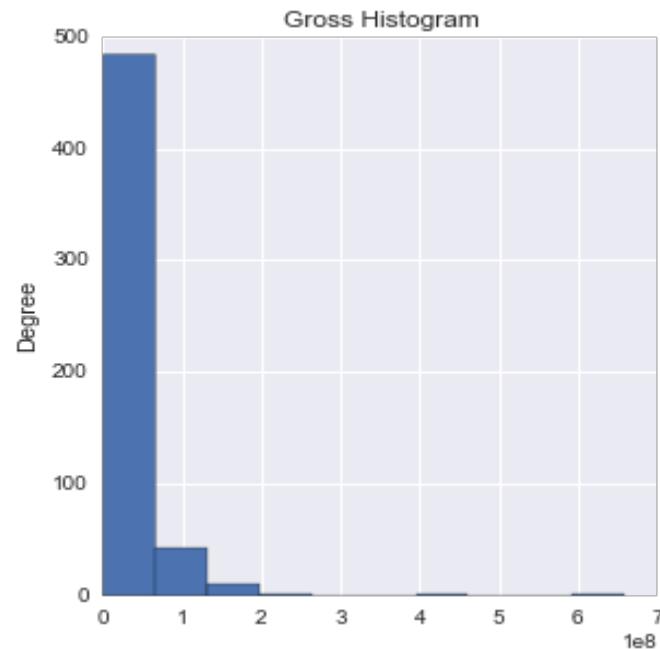
Adjuster:

Traffic

Domestic Total Gross: **\$124,115,725**

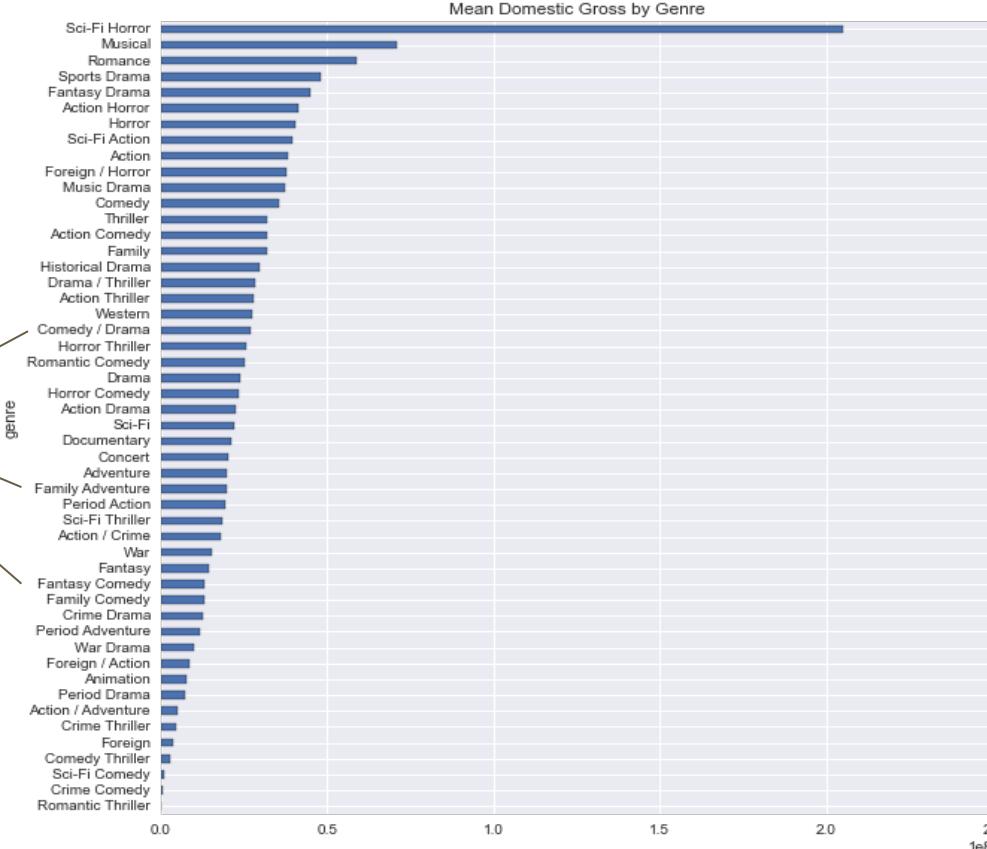
Distributor: USA Films	Release Date: December 27, 2000
Genre: Crime Drama	Runtime: 2 hrs. 27 min.
MPAA Rating: R	Production Budget: N/A

Normalize Gross Data



Categorize Genres into Domestic Gross per Movie

Overlapping
Genres
Examples

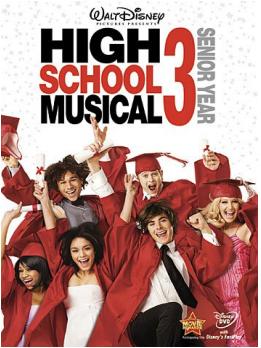


Sci-Fi, Musical, and Romance Movies Make the Highest Gross?

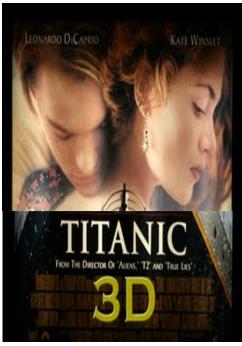
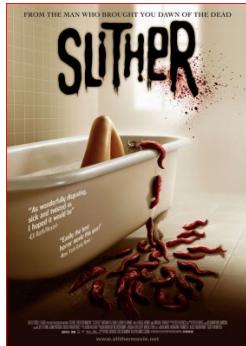
1. Sci Fi Horror



2. Musical

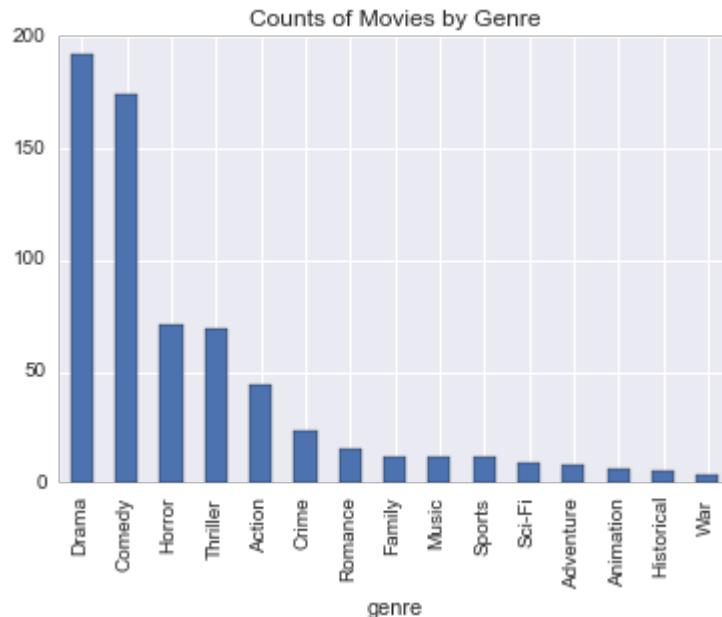


Romance



genre	gross	Count
Sci-Fi Horror	2.051282e+08	2
Musical	7.118108e+07	2
Romance	5.881198e+07	16
Sports Drama	4.825208e+07	12
Fantasy Drama	4.509037e+07	1
Action Horror	4.122693e+07	2
Horror	4.032695e+07	56
Sci-Fi Action	3.945102e+07	2
Action	3.829144e+07	9
Foreign / Horror	3.763462e+07	1
Music Drama	3.751623e+07	10
Comedy	3.536657e+07	103
Thriller	3.211678e+07	32
Action Comedy	3.193628e+07	11
Family	3.190424e+07	8
Historical Drama	2.959049e+07	6
Drama / Thriller	2.825499e+07	15
Action Thriller	2.795851e+07	5
Western	2.731515e+07	4
Comedy / Drama	2.701277e+07	30
Horror Thriller	2.593527e+07	7
Romantic Comedy	2.516095e+07	17
Drama	2.386041e+07	92

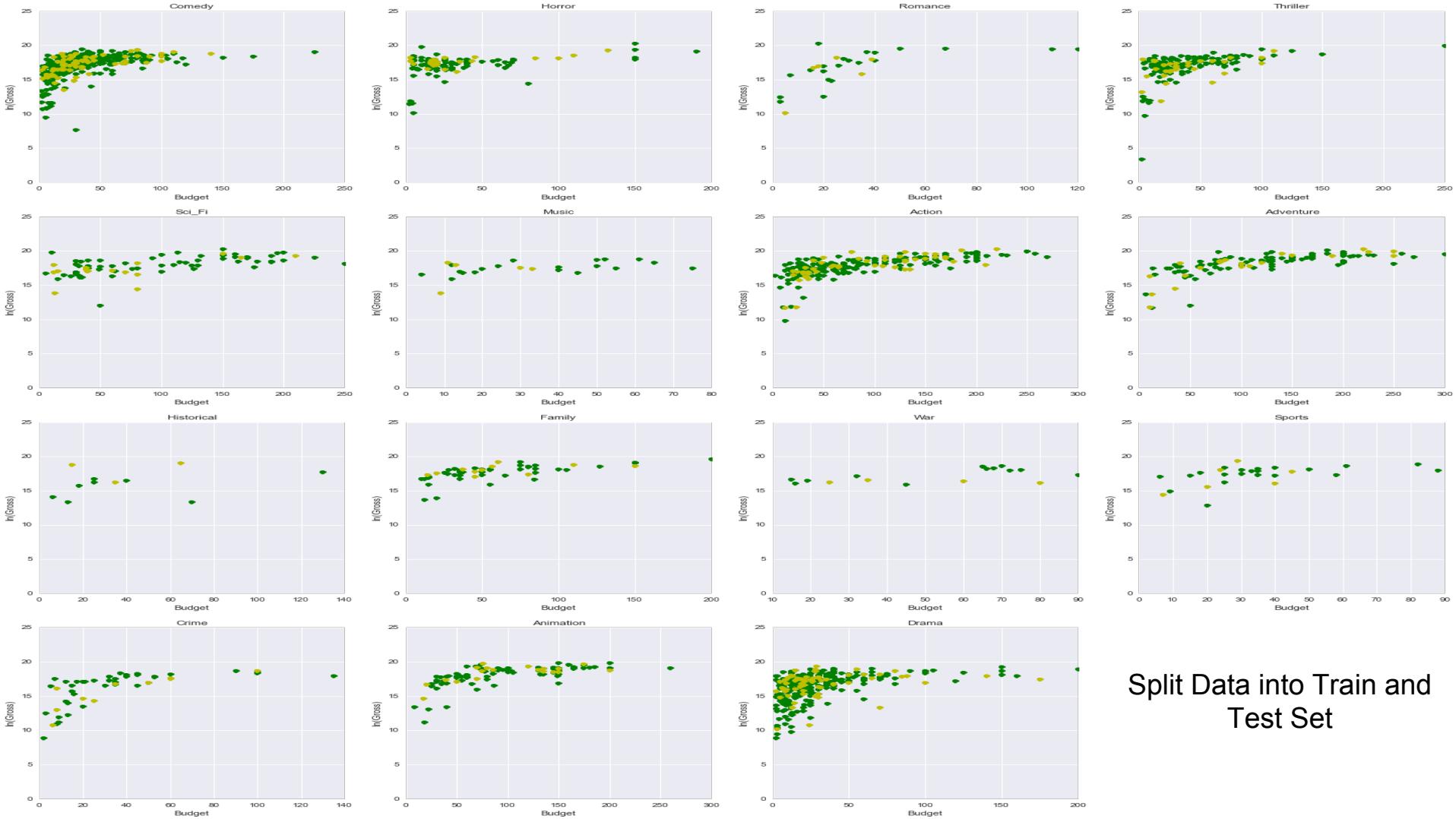
Categorize Movies into Unique Genres



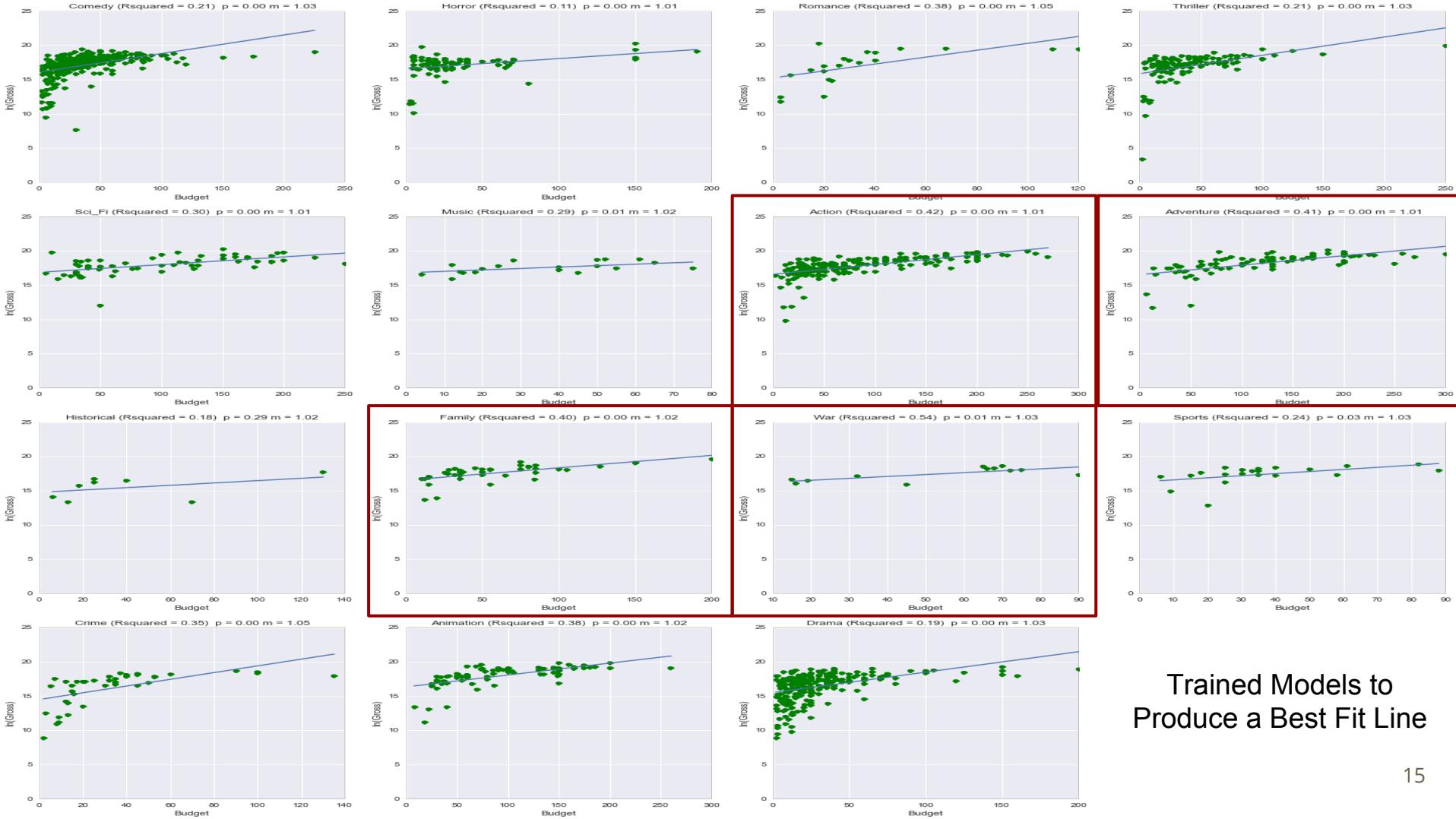
	Average Gross	Count
genre		
Sci-Fi	63104571.666667	9
Romance	58811983.125000	16
Sports	48252081.833333	12
Music	43127042.083333	12
Horror	42822486.408451	71
Comedy	30754760.241379	174
Historical	29590486.333333	6
Action	28083568.386364	44
Thriller	26834460.318841	69
Family	26779763.000000	12
Drama	25751014.437500	192
Adventure	13385511.750000	8
War	12716606.000000	4
Crime	10179030.958333	24
Animation	7841682.428571	7

Analysis and Linear Regression

Given budget and genre, can I predict domestic gross?

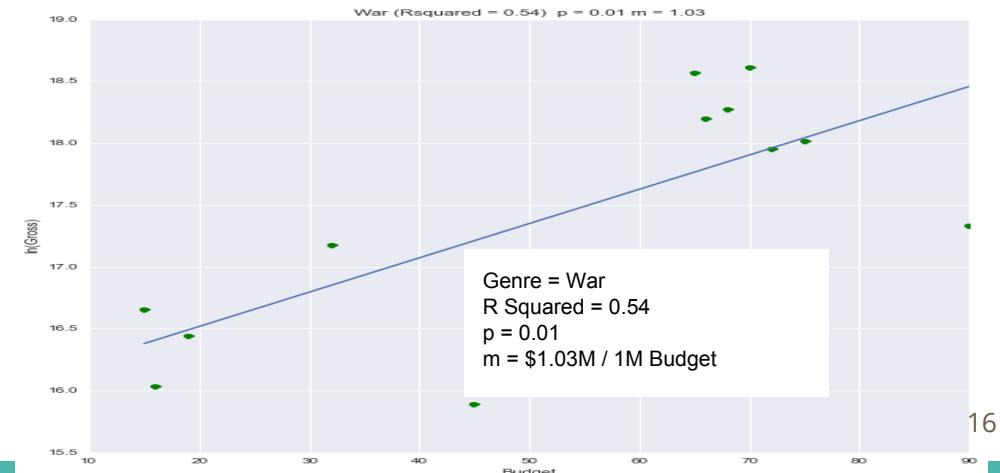
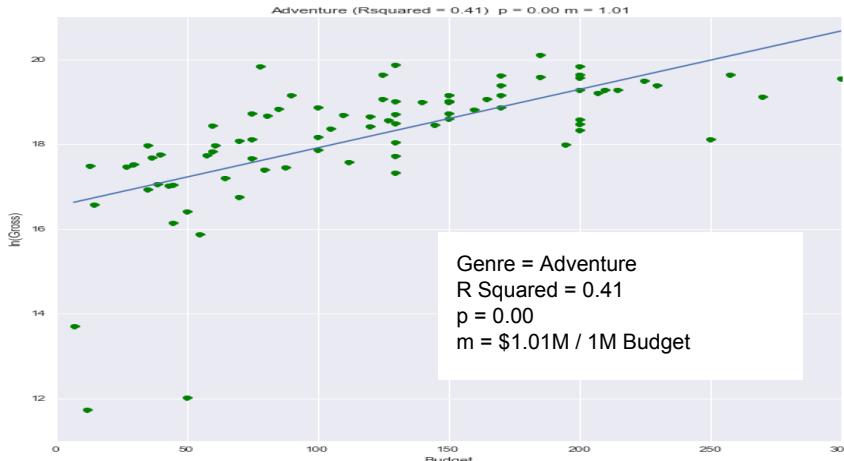
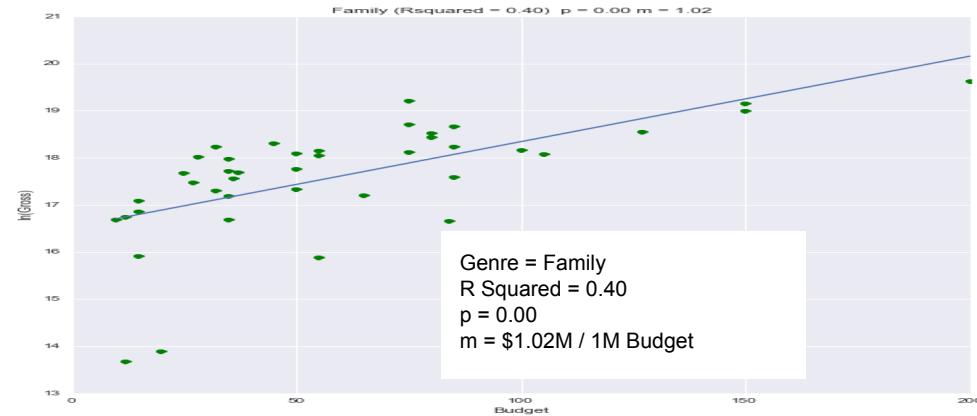
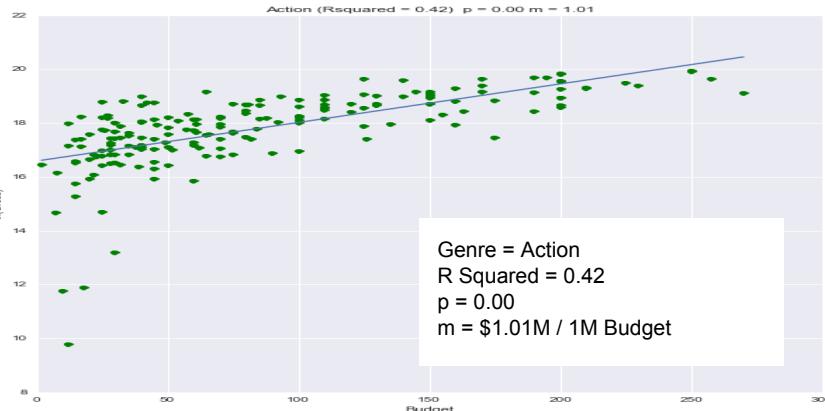


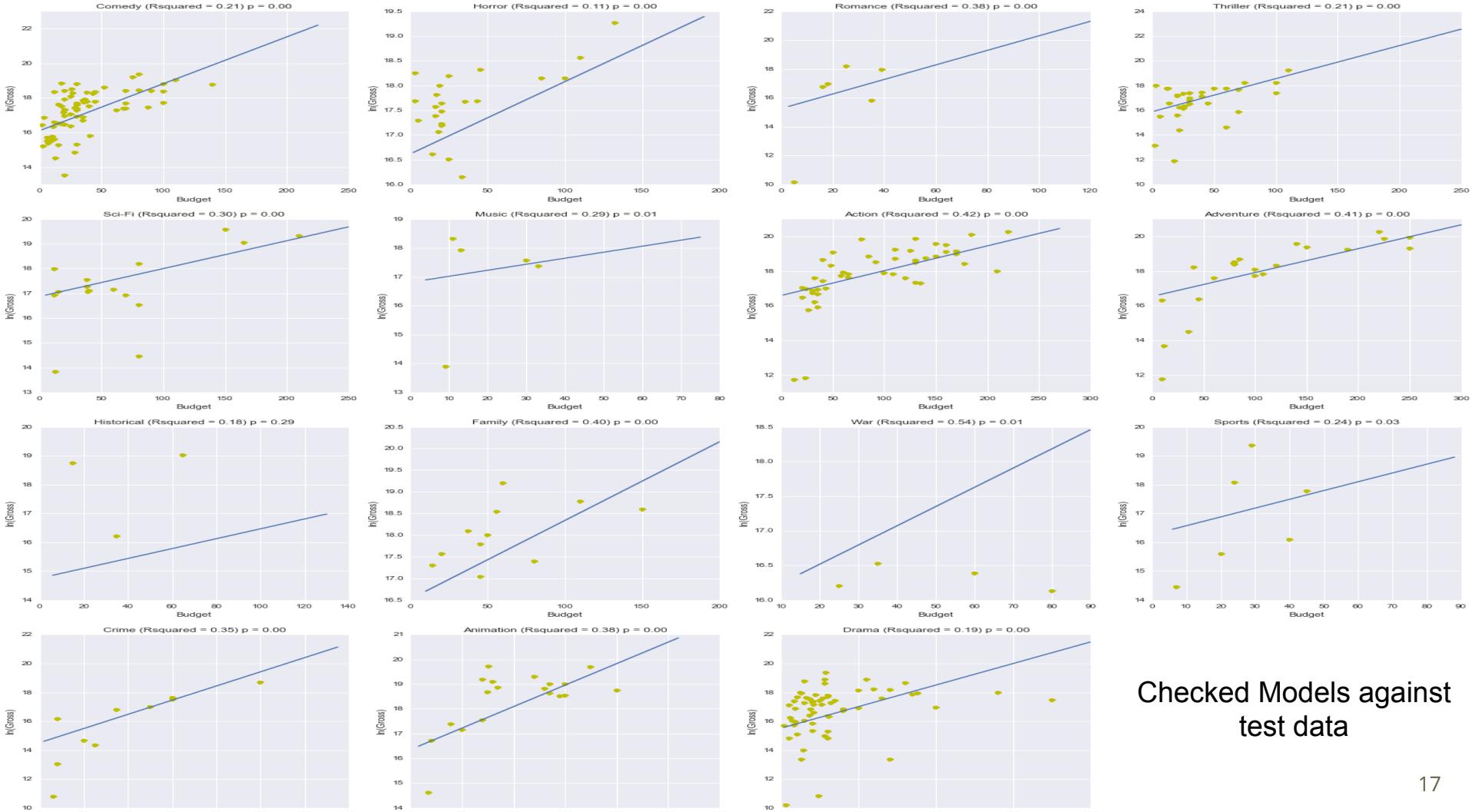
Split Data into Train and
Test Set



Trained Models to
Produce a Best Fit Line

Action, Family, Adventure, and War movies may have better prediction power than other genres when using linear regression.





Checked Models against
test data

Answer: We cannot predict how much a movie will make using Linear regression with high accuracy.

War, Action, Family, and War movies tend to have higher predictive power than other genres.

However, with a smaller budget, you can still make a high grossing movie!



How Can Make Our Model More Awesome?

Use Additional Attributes in Analysis

- Actors
- Directors
- Number Theaters
- Lifetime of Movie

Try a Different Model

**Thank You!
Questions?**