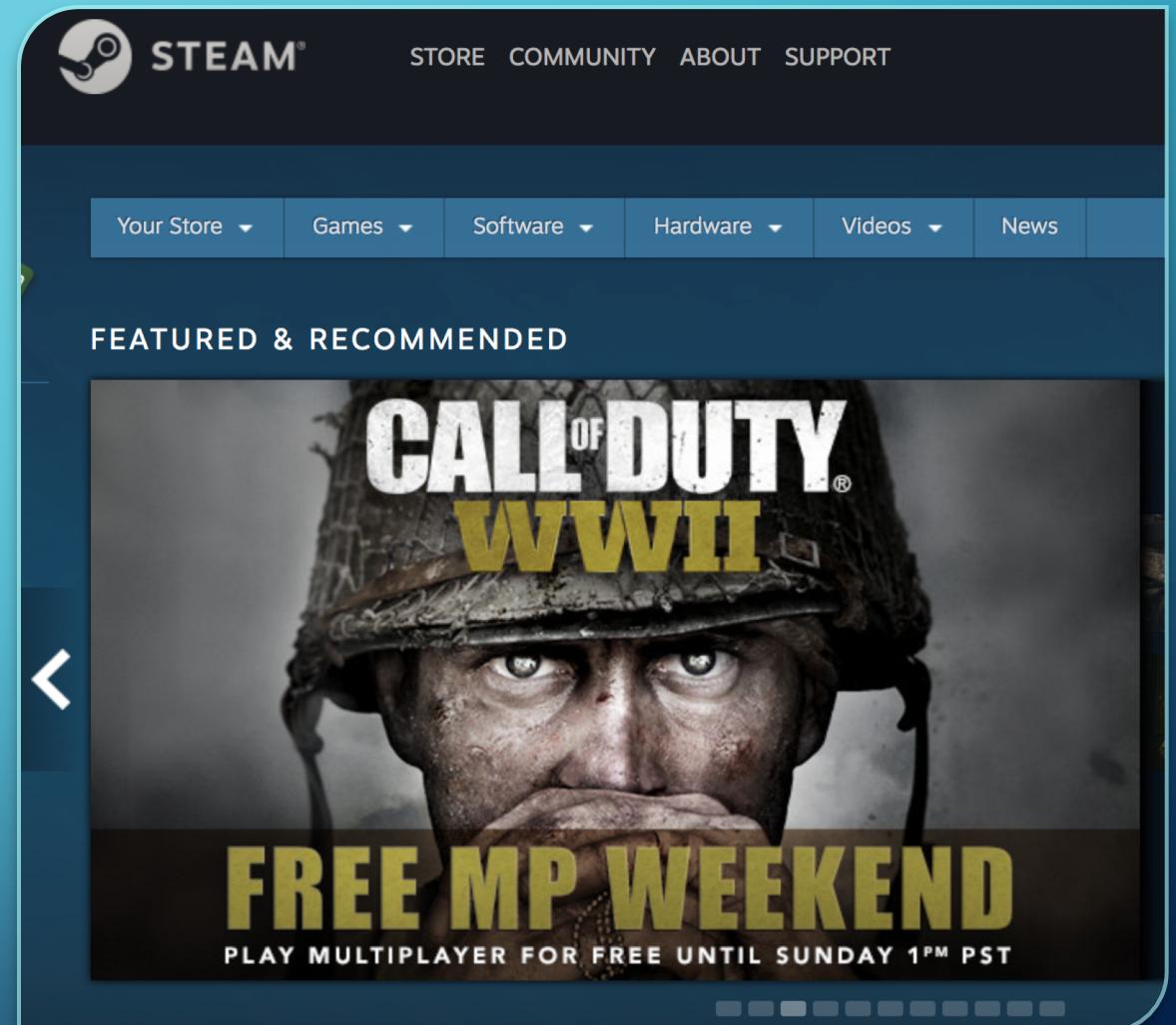


STEAM GAME REVIEWS

MAYA MADHAVAN

RESEARCH QUESTION

- Maximize user engagement with games on Steam => # of reviews
- Steam= largest online game DDP
 - 1.7 million users / month late 2017

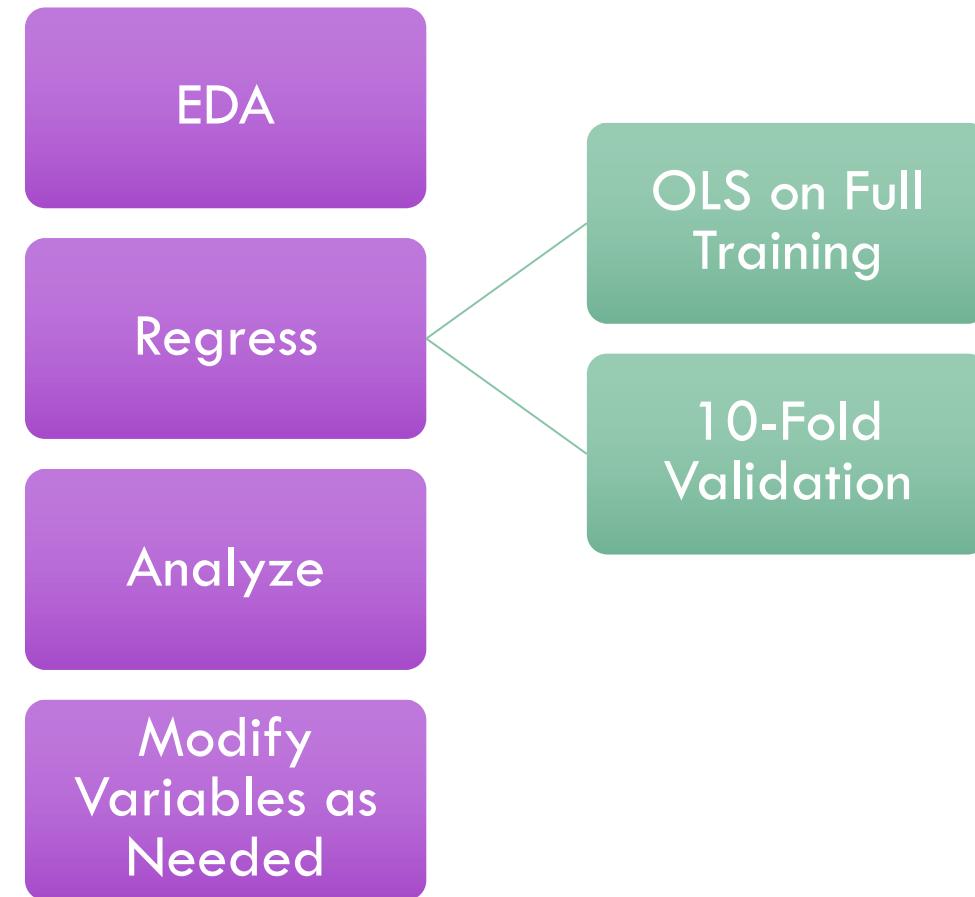


DATASET

- Scrape 10,475 of ~ 49,050 Games
 - Random Arrangement on Steam Search Page
 - => Representative Sample
- Example:

link	system	systems	reviews	release_date	discount	price
title						
Warframe https://store.steampowered.com/app/230410/Warf...	1	['win']	192592	1940	0	0.00
Clicker Heroes 2 https://store.steampowered.com/app/629910/Clic...	2	['win', 'mac']	183	1	0	29.99
Tom Clancy's Rainbow Six® Siege https://store.steampowered.com/app/359550/Tom_...	1	['win']	170518	959	0	14.99
Counter-Strike: Global Offensive https://store.steampowered.com/app/730/Counter...	3	['win', 'mac', 'linux']	2628652	2156	0	14.99
Grand Theft Auto V https://store.steampowered.com/app/271590/Gran...	1	['win']	343780	1191	34	29.99

GENERAL FORMAT OF REGRESSION ANALYSIS



SUCCESSFUL MODIFICATIONS

1

Log Dependent
Variable

2

Drop Windows

3

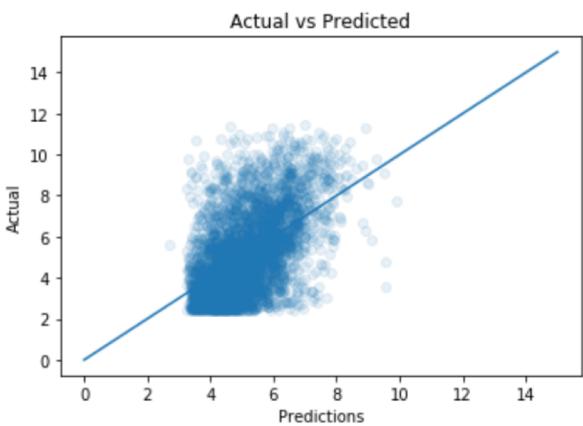
Drop Games
with 0 Reviews

4

Add Polynomial
Features on
Price and
Release Date

BEST MODEL

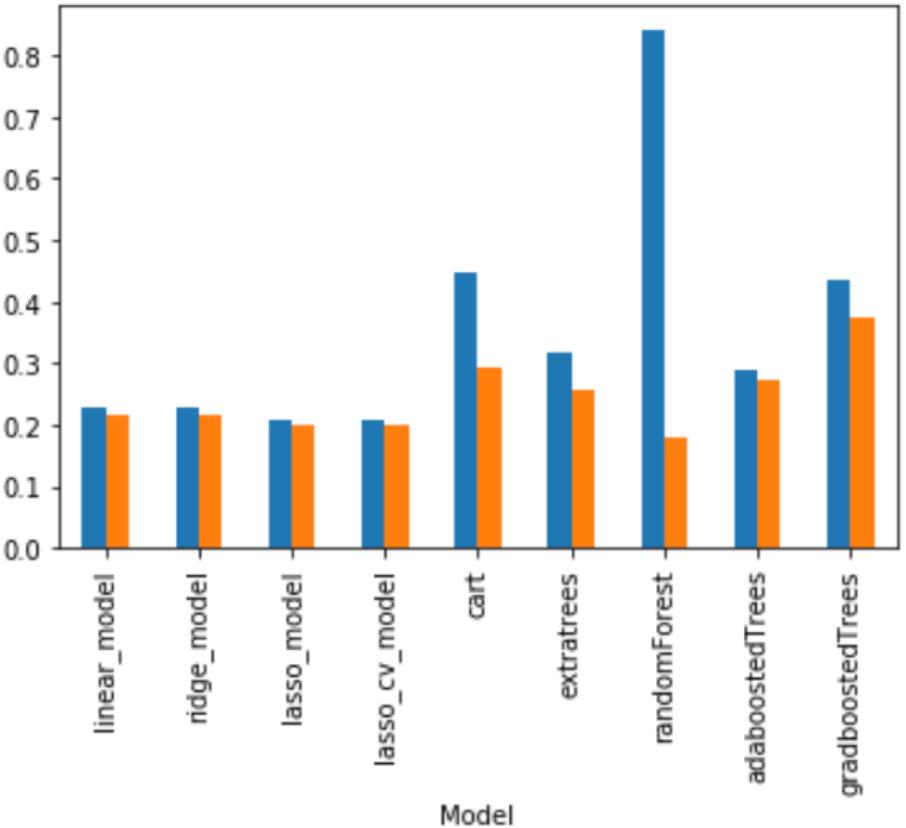
- Polynomial Degree 2
- Logged Y Values
- Dropped Windows and Reviews



OLS Regression Results			
Dep. Variable:	log_reviews	R-squared:	0.270
Model:	OLS	Adj. R-squared:	0.268
Method:	Least Squares	F-statistic:	183.5
Date:	Thu, 19 Jul 2018	Prob (F-statistic):	1.35e-264
Time:	19:45:09	Log-Likelihood:	-7861.9
No. Observations:	3984	AIC:	1.574e+04
Df Residuals:	3975	BIC:	1.580e+04
Df Model:	8		
Covariance Type:	nonrobust		

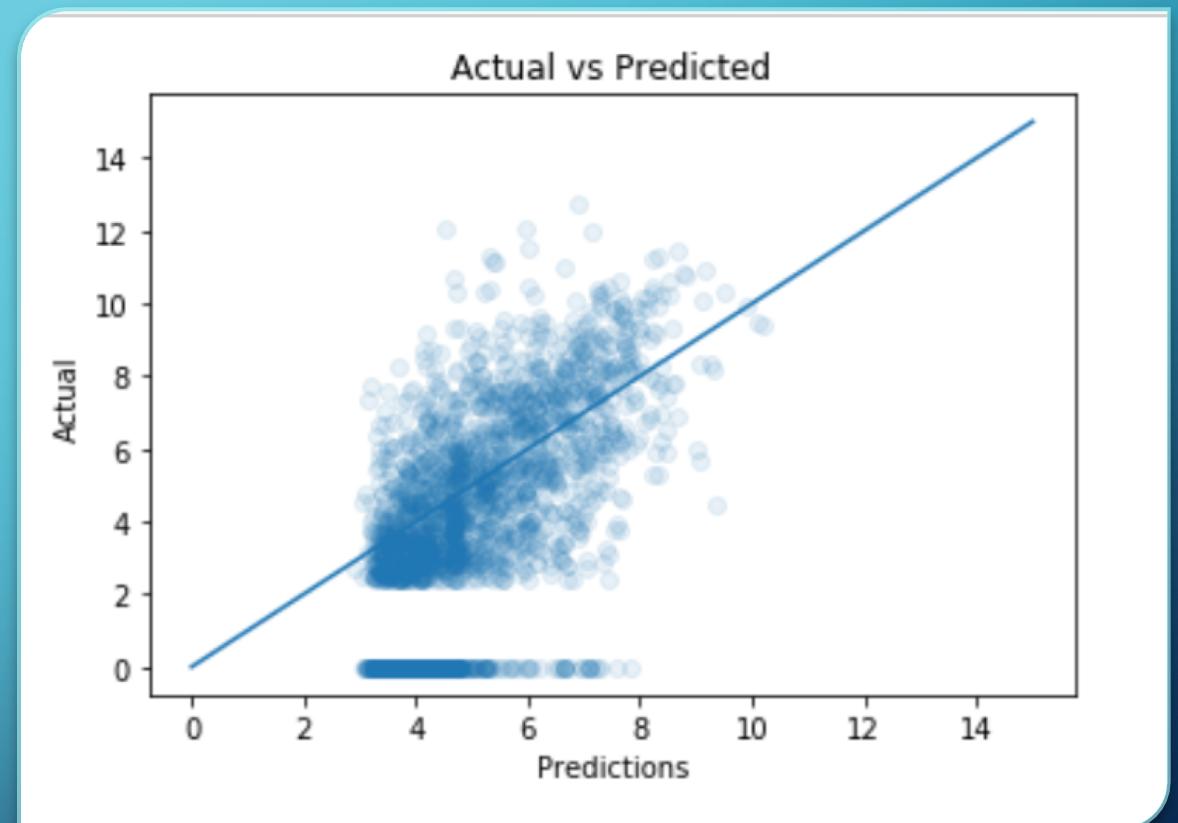
MORE MODELS

- Compare R2 scored
- “Best” Model= Gradient Boosted Trees on Logged Reviews
 - Best in terms of R2, not overfitting



FINAL MODEL + RESULTS

- Final Score on Full Train:
0.4818365496905238
- Final Score on Test:
0.2066513992465474
- **This is a Bad Model**
 - Overfitted
 - High Bias



SO WHAT NOW

- → More Polynomial Features
- → Other Features
- Other Models

title	single	multi	controller_support	indie	action	interface_lang	audio_lang	subtitles_lang
Clicker Heroes 2	1	0		0	1	1	1	1
Counter-Strike: Global Offensive	0	1		1	0	1	25	1
MOTHERGUNSHIP	1	0		1	1	1	7	1
Rocket League®	1	1		1	1	1	12	0
Islands of Nyne: Battle Royale	0	0		0	1	1	1	0
Team Fortress 2	0	1		0	0	1	26	1
Spintires: MudRunner	1	0		1	0	0	9	0
Divinity: Original Sin 2	1	1		1	0	0	7	1
Moonlighter	1	0		1	1	1	10	0

QUESTIONS???