

# What makes a Board Game “Good”?

Predicting eventual ranking on  
BoardGameGeek

Tim Dooley (LASSO)



METIS

# The Board Game Industry

“Over \$12 billion by 2023” - [Industry Forecast](#)



“In 2019, fans pledged more than \$176 million [on Kickstarter] toward tabletop games — up 6.8% over the previous year”

- [Guardian Article, 5 July 2020](#)

# How can a Board Game Publisher/Designer expect their new Board Game to perform?

BoardGameGeek is a collection of thousands of Board Games

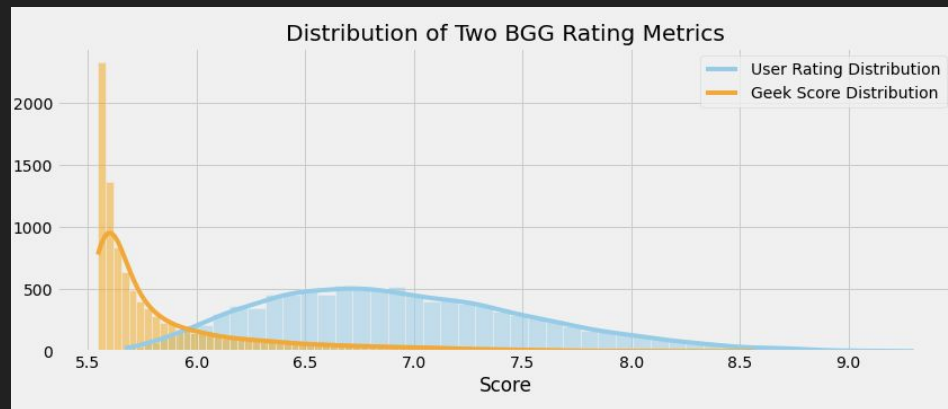
So much data, but ranking is not commensurate with user reviews

Objectives:

Scrape 10,000 games from BGG

Build a Supervised Learning Model

to predict game ranking



# Methodology

## Collection

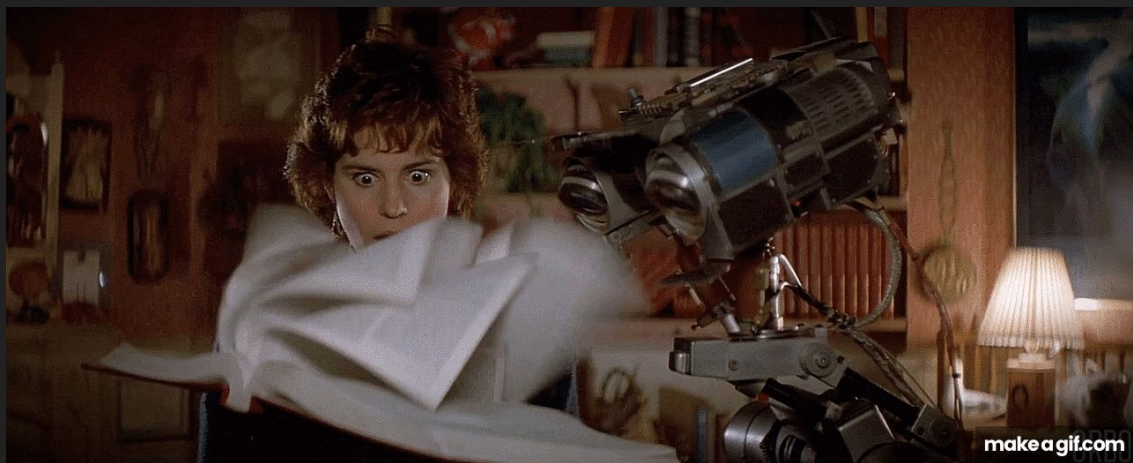
Requests/BeautifulSoup

Selenium (dynamically coded database)

## Modelling

SKLearn and StatsModels for Linear Regression

Imputation with MICE for NaN values that couldn't be scraped



Data columns (total 108 columns):

#	Column	Non-Null Count		Dtype
0	link	10019	non-null	object
1	rank	10019	non-null	float64
2	desc	10019	non-null	object
3	geek	10019	non-null	float64
4	rating	10019	non-null	float64
5	num_voters	10019	non-null	int64
6	price	10019	non-null	object
7	average	10019	non-null	object
8	avgweight	10019	non-null	float64
9	baverage	10019	non-null	float64
10	maxplayers	10018	non-null	float64
11	maxplaytime	10018	non-null	float64
12	minage	10018	non-null	float64
13	minplayers	10018	non-null	float64
14	minplaytime	10018	non-null	float64
15	numcomments	10019	non-null	float64
16	numfans	10019	non-null	float64
17	numgeeklists	10019	non-null	float64
18	numhasparts	10019	non-null	float64
19	numowned	10019	non-null	float64
20	numplays	10019	non-null	float64
21	numplays_month	10019	non-null	float64
22	numprevowned	10019	non-null	float64
23	numtrading	10019	non-null	float64
24	numwanting	10019	non-null	float64
25	numwantparts	10019	non-null	float64
26	numweights	10019	non-null	float64
27	numwish	10019	non-null	float64
28	numwishlistcomments	10019	non-null	float64
29	playmonth	10019	non-null	object
30	price_mkt	3332	non-null	float64
31	stddev	10019	non-null	float64
32	usersrated	10019	non-null	float64
33	views	10019	non-null	float64
34	walmart_price	260	non-null	object
35	yearpublished	10018	non-null	object
36	genres	10010	non-null	object
37	designer	9906	non-null	object
38	publish	10008	non-null	object
39	strategy_war	10019	non-null	int64
40	strategy	10019	non-null	int64
41	family	10019	non-null	int64
42	thematic	10019	non-null	int64
43	war	10019	non-null	int64
44	party	10019	non-null	int64
45	abstract	10019	non-null	int64
46	rank_des	9906	non-null	float64
47	geek_des	9906	non-null	float64
48	rating_des	9906	non-null	float64
49	num_voters_des	9906	non-null	float64
50	baverage_des	9906	non-null	float64

51	maxplayers_des	9906	non-null	float64
52	maxplaytime_des	9906	non-null	float64
53	minage_des	9906	non-null	float64
54	minplayers_des	9906	non-null	float64
55	minplaytime_des	9906	non-null	float64
56	numcomments_des	9906	non-null	float64
57	numfans_des	9906	non-null	float64
58	numgeeklists_des	9906	non-null	float64
59	numowned_des	9906	non-null	float64
60	numplays_des	9906	non-null	float64
61	numplays_month_des	9906	non-null	float64
62	numtrading_des	9906	non-null	float64
63	numwanting_des	9906	non-null	float64
64	numwantparts_des	9906	non-null	float64
65	numweights_des	9906	non-null	float64
66	numwish_des	9906	non-null	float64
67	price_mkt_des	7238	non-null	float64
68	views_des	9906	non-null	float64
69	strategy_war_des	9906	non-null	float64
70	strategy_des	9906	non-null	float64
71	family_des	9906	non-null	float64
72	thematic_des	9906	non-null	float64
73	war_des	9906	non-null	float64
74	party_des	9906	non-null	float64
75	abstract_des	9906	non-null	float64
76	numdes	9906	non-null	float64
77	rank_pub	10008	non-null	float64
78	geek_pub	10008	non-null	float64
79	rating_pub	10008	non-null	float64
80	num_voters_pub	10008	non-null	float64
81	baverage_pub	10008	non-null	float64
82	maxplayers_pub	10008	non-null	float64
83	maxplaytime_pub	10008	non-null	float64
84	minage_pub	10008	non-null	float64
85	minplayers_pub	10008	non-null	float64
86	minplaytime_pub	10008	non-null	float64
87	numcomments_pub	10008	non-null	float64
88	numfans_pub	10008	non-null	float64
89	numgeeklists_pub	10008	non-null	float64
90	numowned_pub	10008	non-null	float64
91	numplays_pub	10008	non-null	float64
92	numplays_month_pub	10008	non-null	float64
93	numtrading_pub	10008	non-null	float64
94	numwanting_pub	10008	non-null	float64
95	numwantparts_pub	10008	non-null	float64
96	numweights_pub	10008	non-null	float64
97	numwish_pub	10008	non-null	float64
98	price_mkt_pub	8604	non-null	float64
99	views_pub	10008	non-null	float64
100	strategy_war_pub	10008	non-null	float64
101	strategy_pub	10008	non-null	float64
102	family_pub	10008	non-null	float64
103	thematic_pub	10008	non-null	float64
104	war_pub	10008	non-null	float64
105	party_pub	10008	non-null	float64
106	abstract_pub	10008	non-null	float64
107	numpub	10008	non-null	float64

108	maxplayers_des	9906	non-null	float64
109	maxplaytime_des	9906	non-null	float64
110	minage_des	9906	non-null	float64
111	minplayers_des	9906	non-null	float64
112	minplaytime_des	9906	non-null	float64
113	numcomments_des	9906	non-null	float64
114	numfans_des	9906	non-null	float64
115	numgeeklists_des	9906	non-null	float64
116	numowned_des	9906	non-null	float64
117	numplays_des	9906	non-null	float64
118	numplays_month_des	9906	non-null	float64
119	numtrading_des	9906	non-null	float64
120	numwanting_des	9906	non-null	float64
121	numwantparts_des	9906	non-null	float64
122	numweights_des	9906	non-null	float64
123	numwish_des	9906	non-null	float64
124	price_mkt_des	7238	non-null	float64
125	views_des	9906	non-null	float64
126	strategy_war_des	9906	non-null	float64
127	strategy_des	9906	non-null	float64
128	family_des	9906	non-null	float64
129	thematic_des	9906	non-null	float64
130	war_des	9906	non-null	float64
131	party_des	9906	non-null	float64
132	abstract_des	9906	non-null	float64
133	numdes	9906	non-null	float64
134	rank_pub	10008	non-null	float64
135	geek_pub	10008	non-null	float64
136	rating_pub	10008	non-null	float64
137	num_voters_pub	10008	non-null	float64
138	baverage_pub	10008	non-null	float64
139	maxplayers_pub	10008	non-null	float64
140	maxplaytime_pub	10008	non-null	float64
141	minage_pub	10008	non-null	float64
142	minplayers_pub	10008	non-null	float64
143	minplaytime_pub	10008	non-null	float64
144	numcomments_pub	10008	non-null	float64
145	numfans_pub	10008	non-null	float64
146	numgeeklists_pub	10008	non-null	float64
147	numowned_pub	10008	non-null	float64
148	numplays_pub	10008	non-null	float64
149	numplays_month_pub	10008	non-null	float64
150	numtrading_pub	10008	non-null	float64
151	numwanting_pub	10008	non-null	float64
152	numwantparts_pub	10008	non-null	float64
153	numweights_pub	10008	non-null	float64
154	numwish_pub	10008	non-null	float64
155	price_mkt_pub	8604	non-null	float64
156	views_pub	10008	non-null	float64
157	strategy_war_pub	10008	non-null	float64
158	strategy_pub	10008	non-null	float64
159	family_pub	10008	non-null	float64
160	thematic_pub	10008	non-null	float64
161	war_pub	10008	non-null	float64
162	party_pub	10008	non-null	float64
163	abstract_pub	10008	non-null	float64
164	numpub	10008	non-null	float64



# Methodology: Features

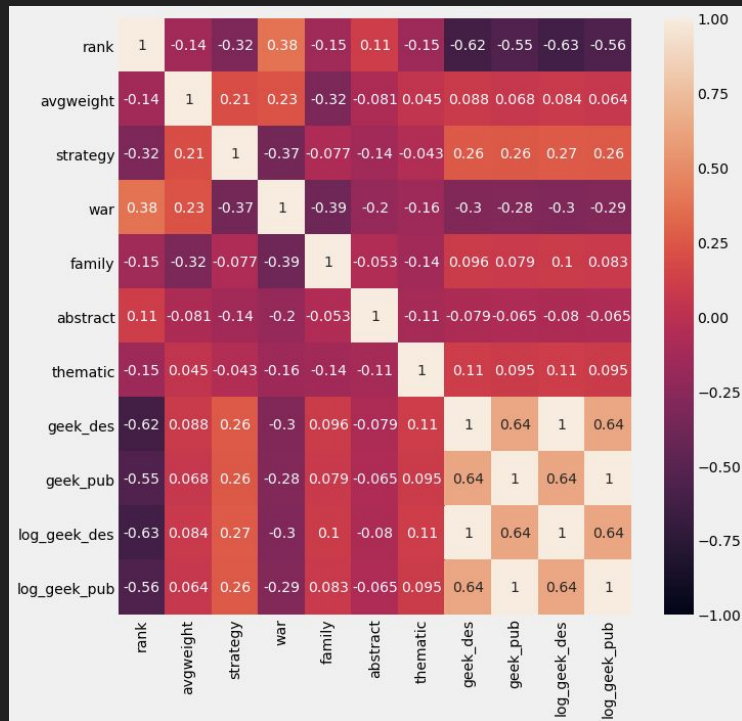
Average Weight, aka Complexity

Genres (dummy encoded)

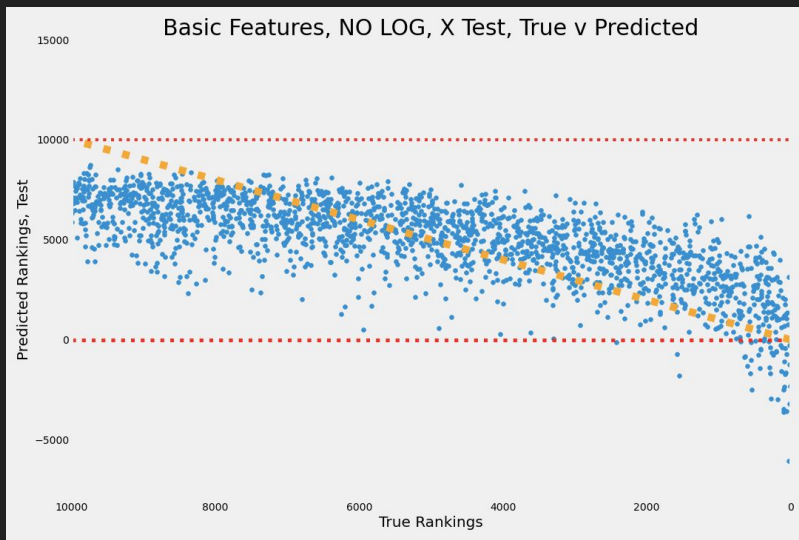
The Geek Score of each Designer and Publisher

Importantly: Log transforms of those scores and rank

Avoid 'leakage' and assume only basic information about game





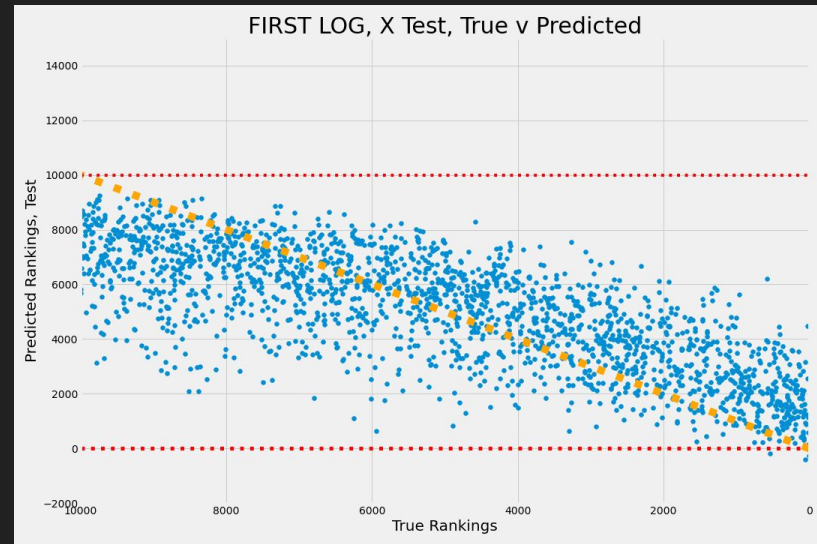


<  $R^2: 0.48$

$R^2: 0.56$  >

This model curves toward the higher true rankings

Tends to produce negative predicted rankings for “good” games



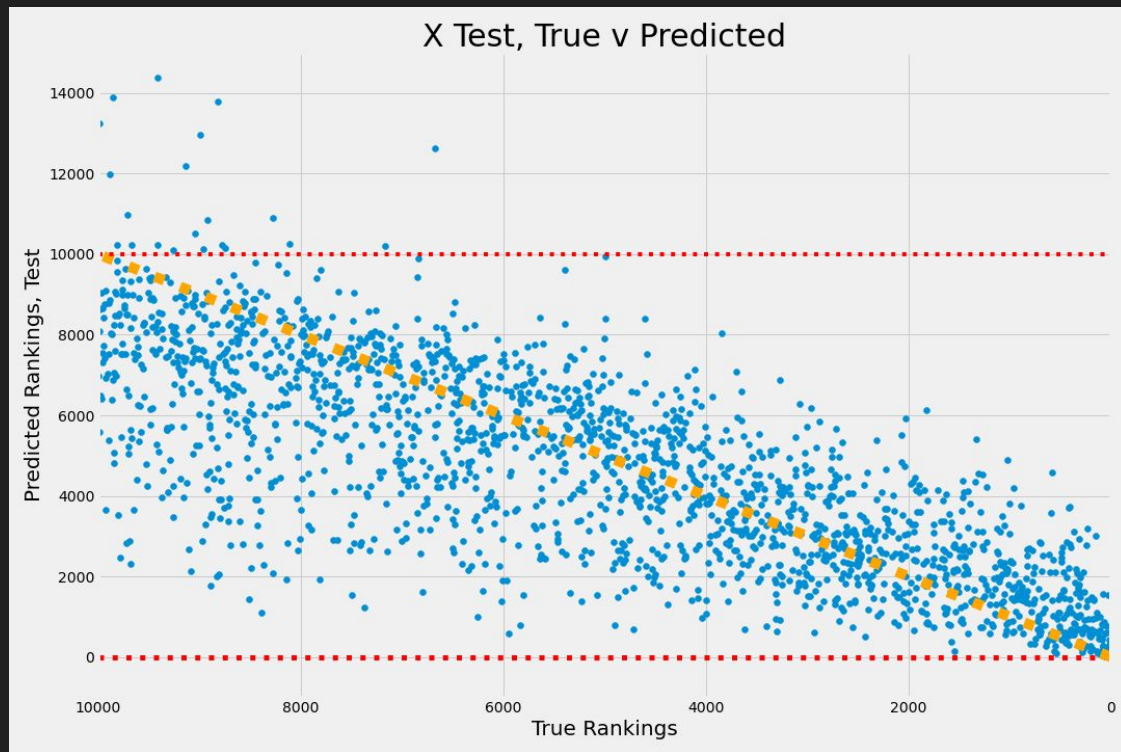
Upon applying log transforms to Designer and Publisher Geek scores, the model improves slightly

# Results: Polyfit/Log Transforms

Final R2 Train: 0.60

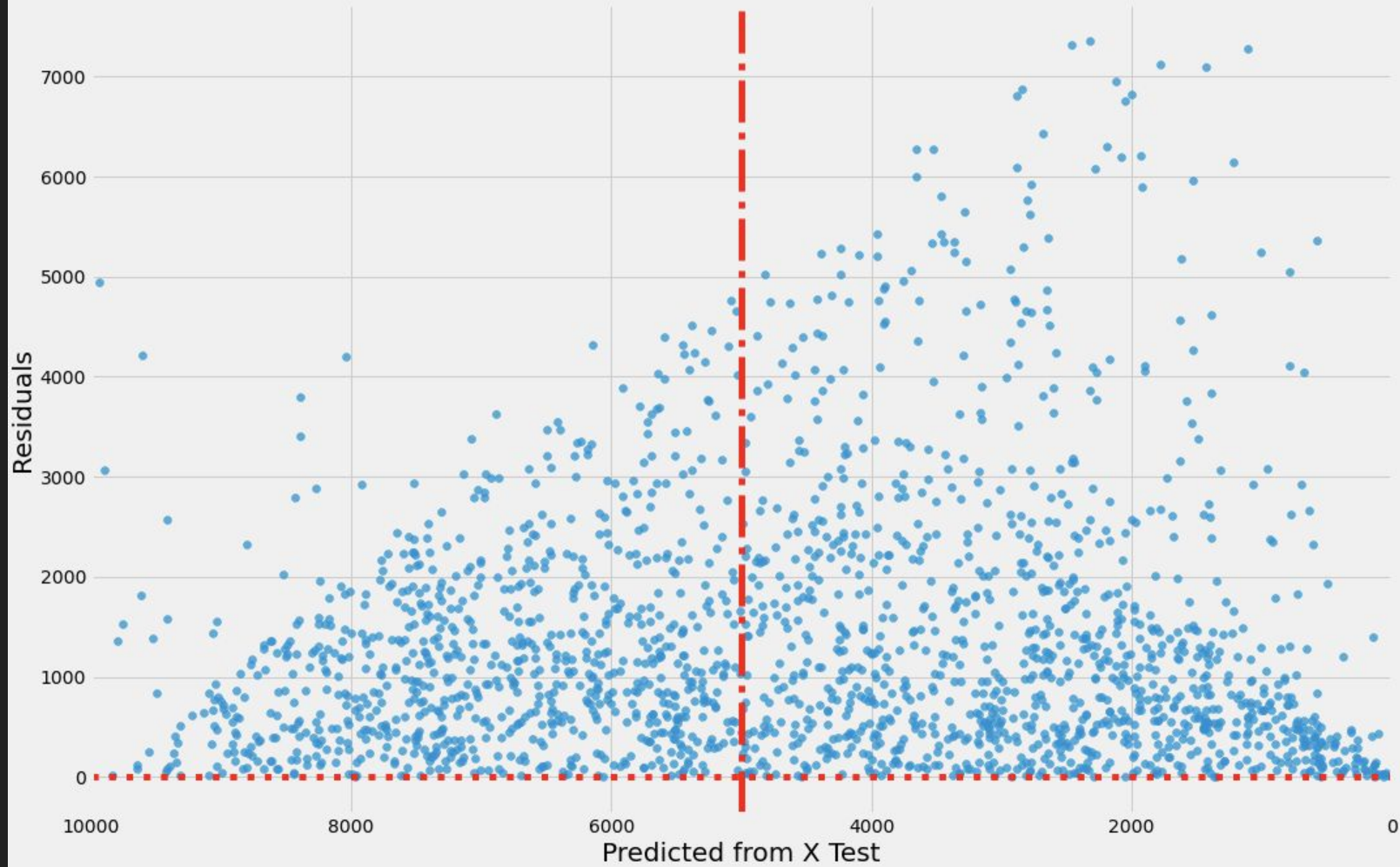
Final R2 Test: 0.64

R2 Ratio: 0.943





# Residual Test Data



# What can you do to improve your rank?

**Top 1/3!** Produce **complex** games that involve **strategy** and **thematic** elements (think fantasy, sci-fi) that are accessible for the whole **family**

Hire a decent **designer**:

The mean average designer: 5.871 Geek

Hire 5.971: see your game's rank rise 500+ places!

**Avoid abstract** concepts (don't reinvent Chess!)

**Avoid war** games

# Future Considerations

Of course, there's more to "good" than rank.

- Industry playtest data

- Better price data

- Problem of 'collector items'

# Thank You

