NEW YORK TIMES BESTSELLER ANALYSIS

Presented by **Linh Mai**

October 6th, 2020

objectives

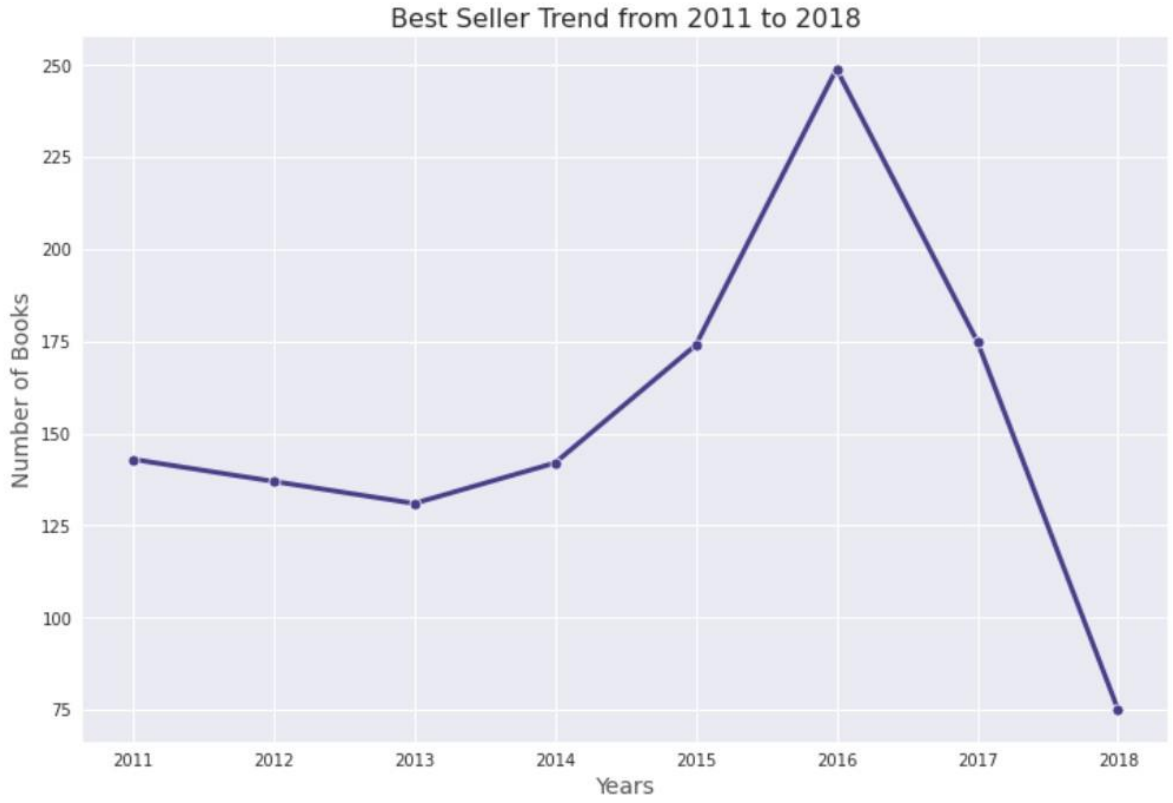
- Popularity, exposure, guideline
- Generally increase sale by ~14%

  Debut authors: increase the sale by >50%

- Best seller trend decreases since 2016

  (Data only for the first half of the year 2018)

The New York Times



overview

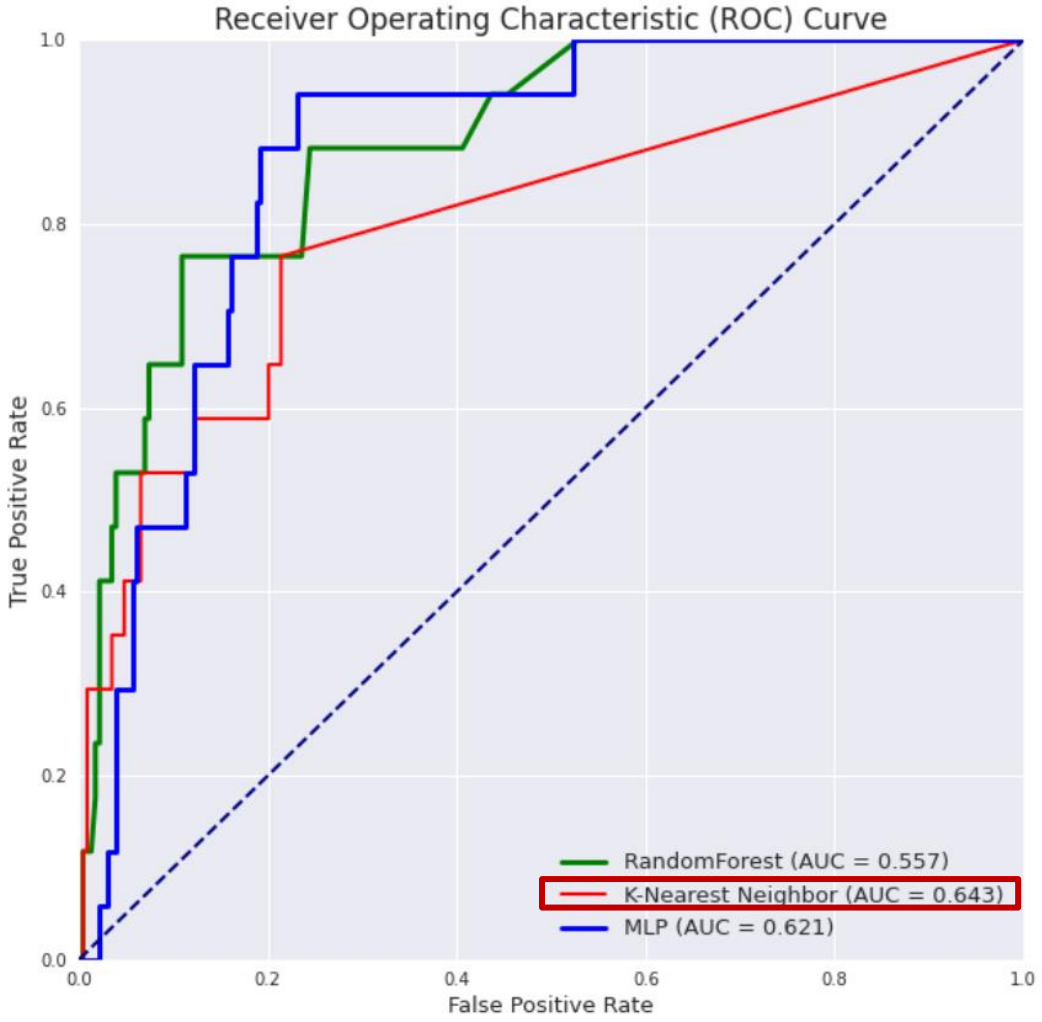
New York Times Bestsellers
from Feb-2011 to Jun-2018

Book review statistics based on
the ISBNs list from **data.world**

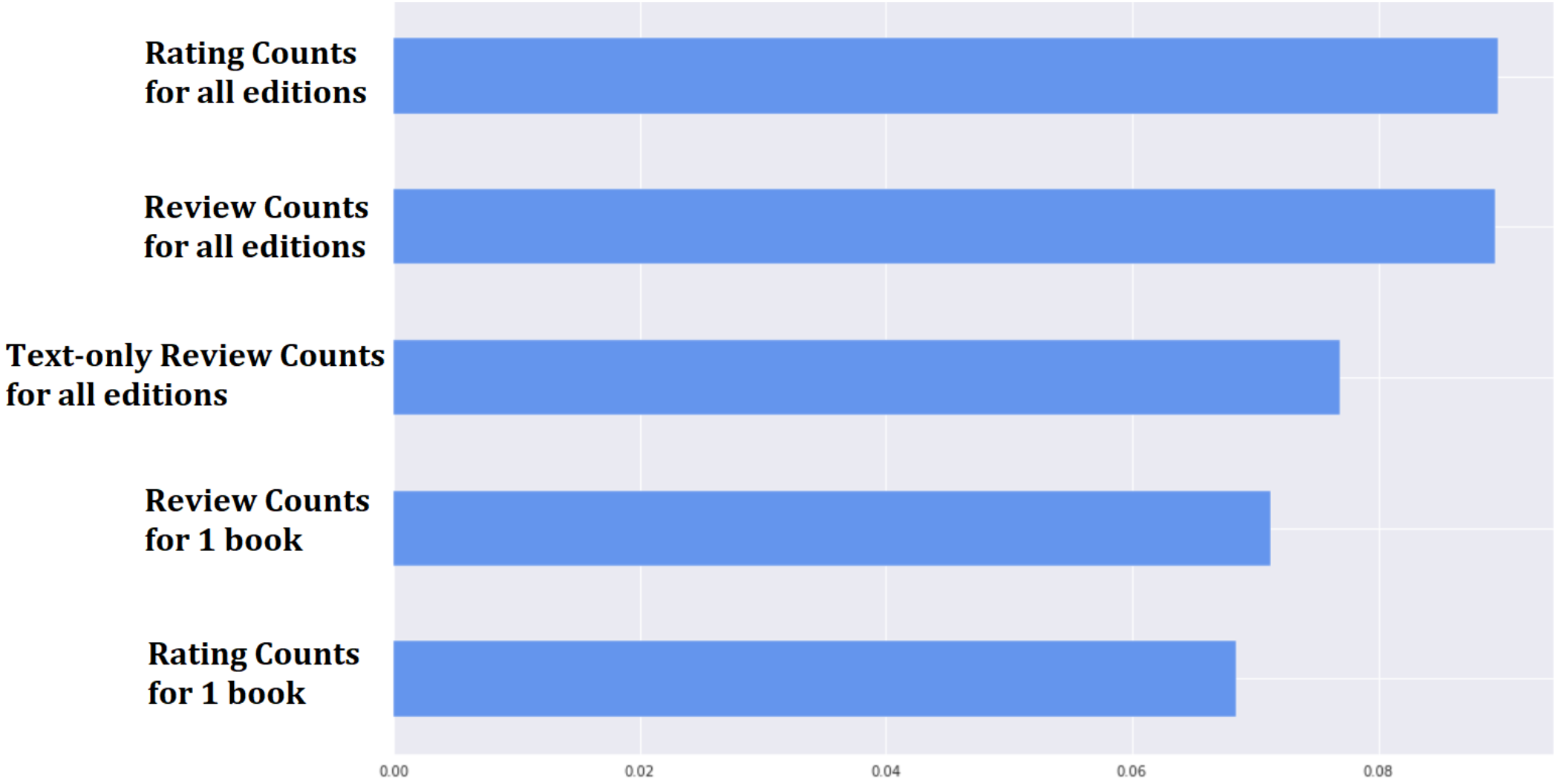
- **Raw data**: 2248 uncleaned data points
- **Cleaned & combined data**:
  1226 data points with 10 features
  2216 data points with 3 features

top 3 models' performance

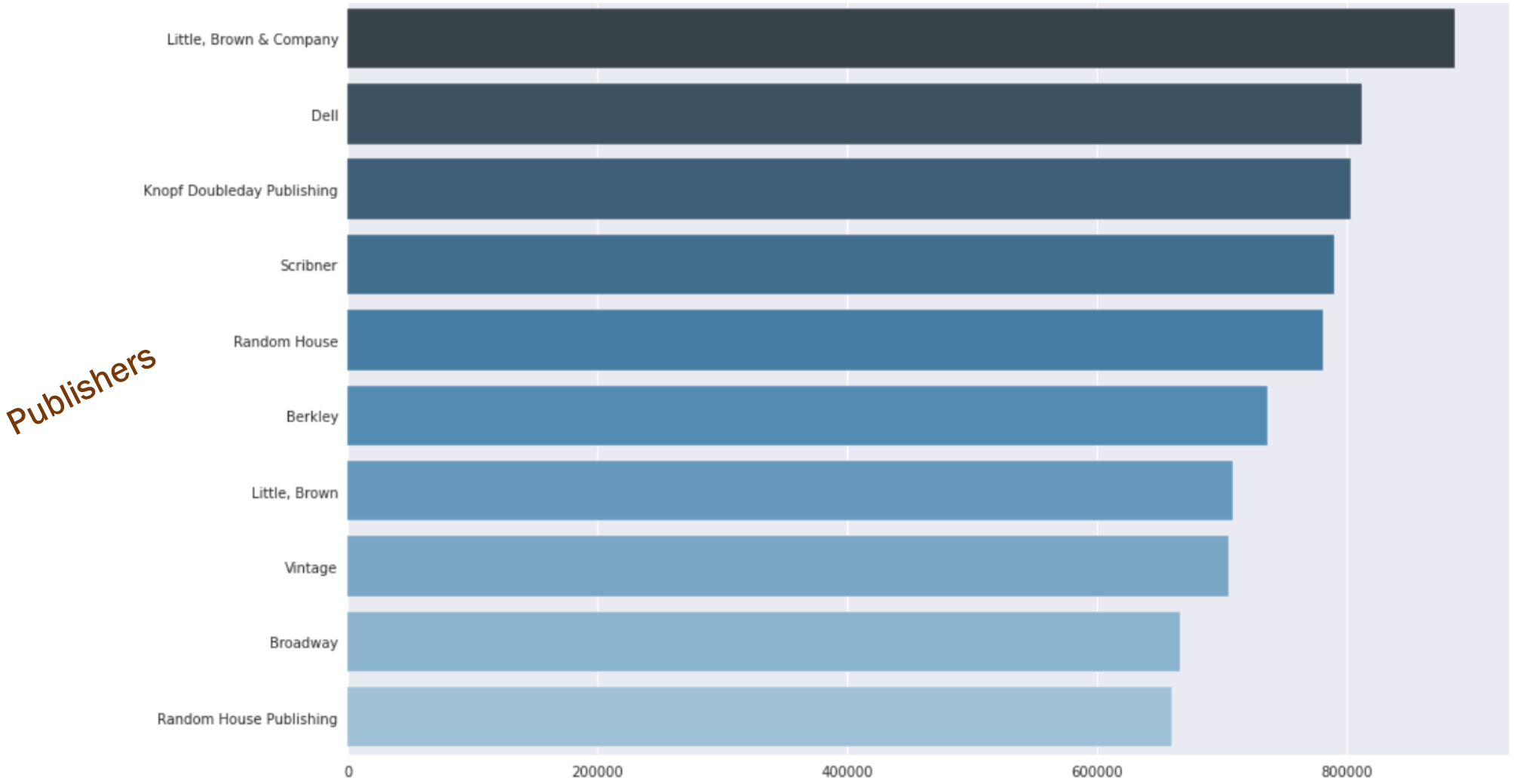
	Precision	AUC
K-Nearest Neighbor	0.71	0.64
Random Forest	*0.67*	0.56
MLP	0.28	*0.62*



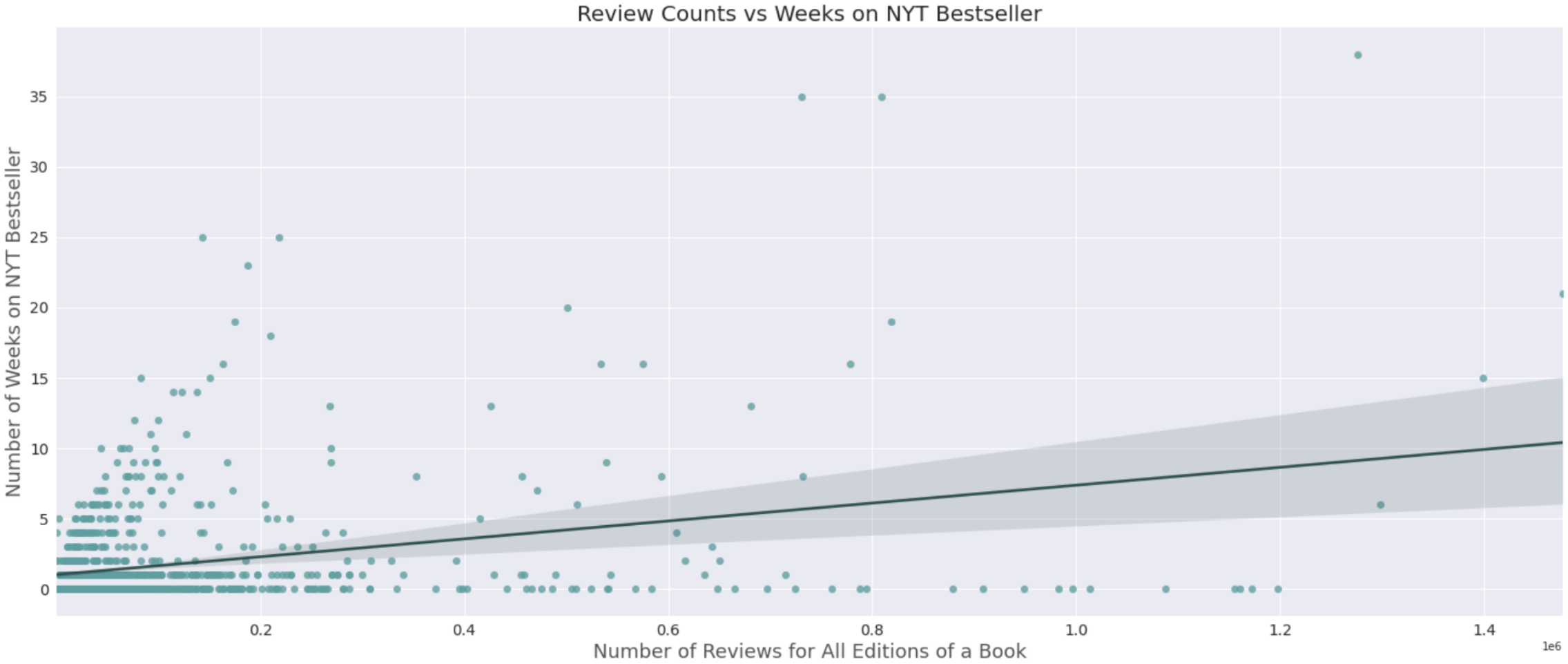
important features



recommendations

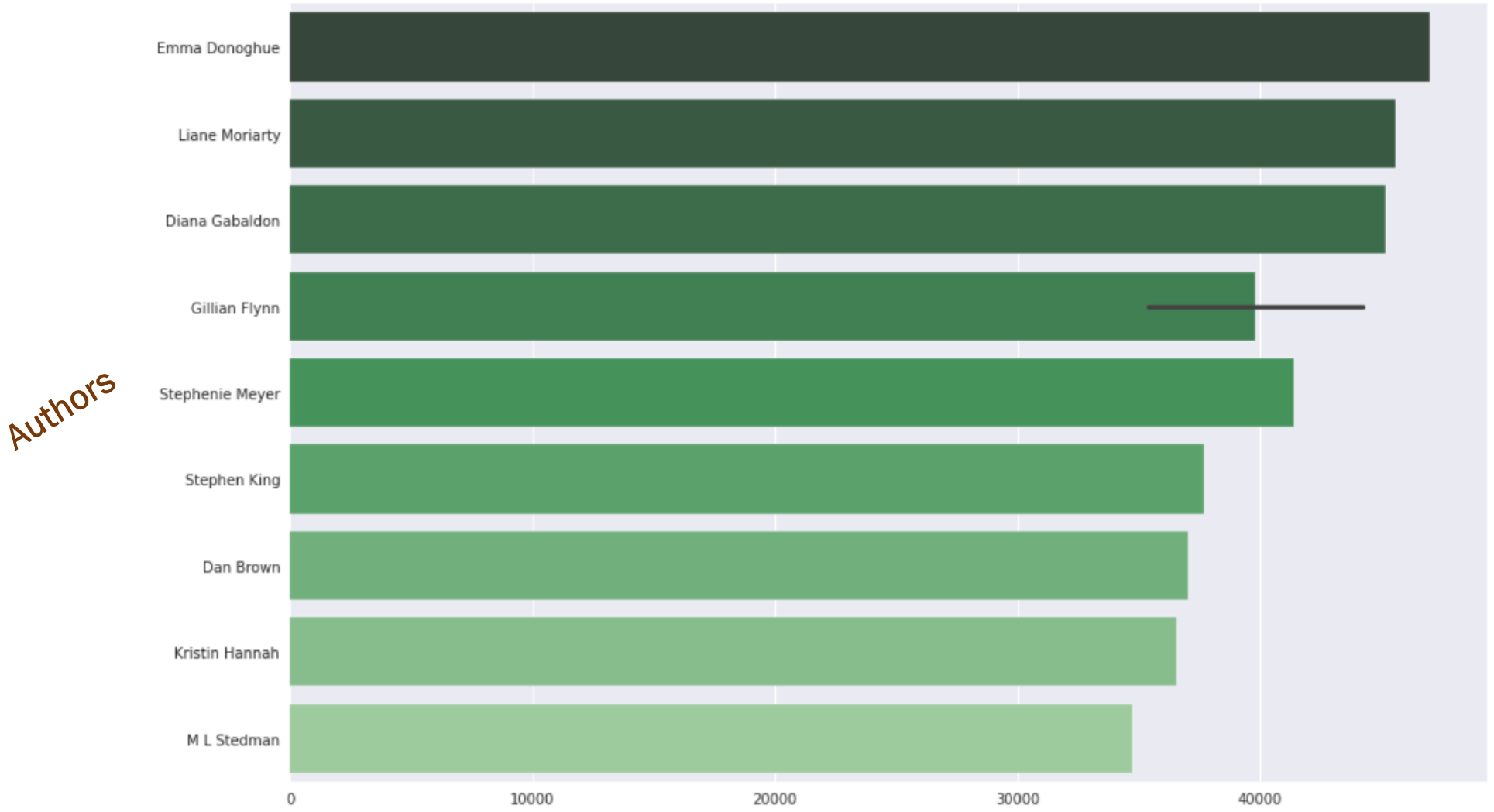


recommendations



Average: 16 weeks of staying on NYT Bestseller

recommendations



Number of books

1
3
6
3
2
11
2
8
1

Average = 4

recommendations

Review Counts for all Editions & Titles

Number of Words in Title

future considerations

More data from other existing datasets or scrapping webpages Google Books, Amazon

In-depth research on top features by using different type of modeling: regression

THANK YOU