

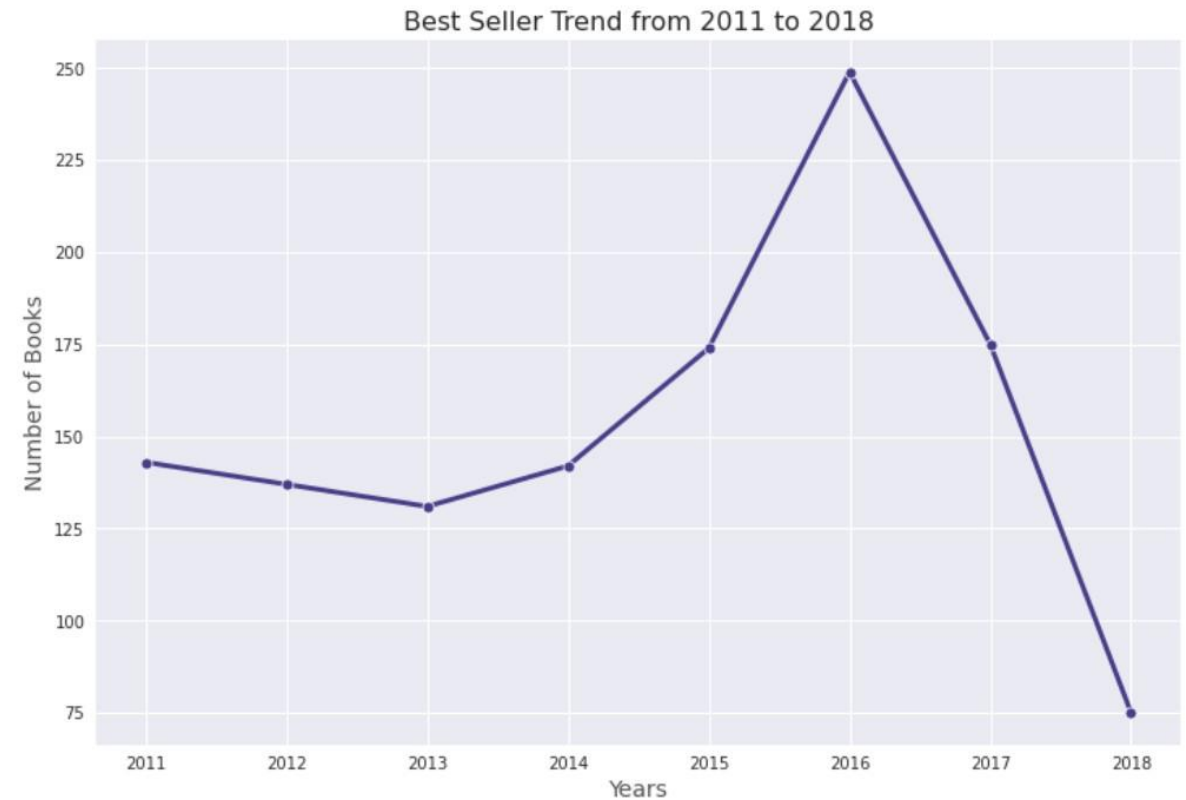
NEW YORK TIMES BESTSELLER ANALYSIS

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

The New York Times



overview



New York Times Bestsellers
from 2011 to 2018



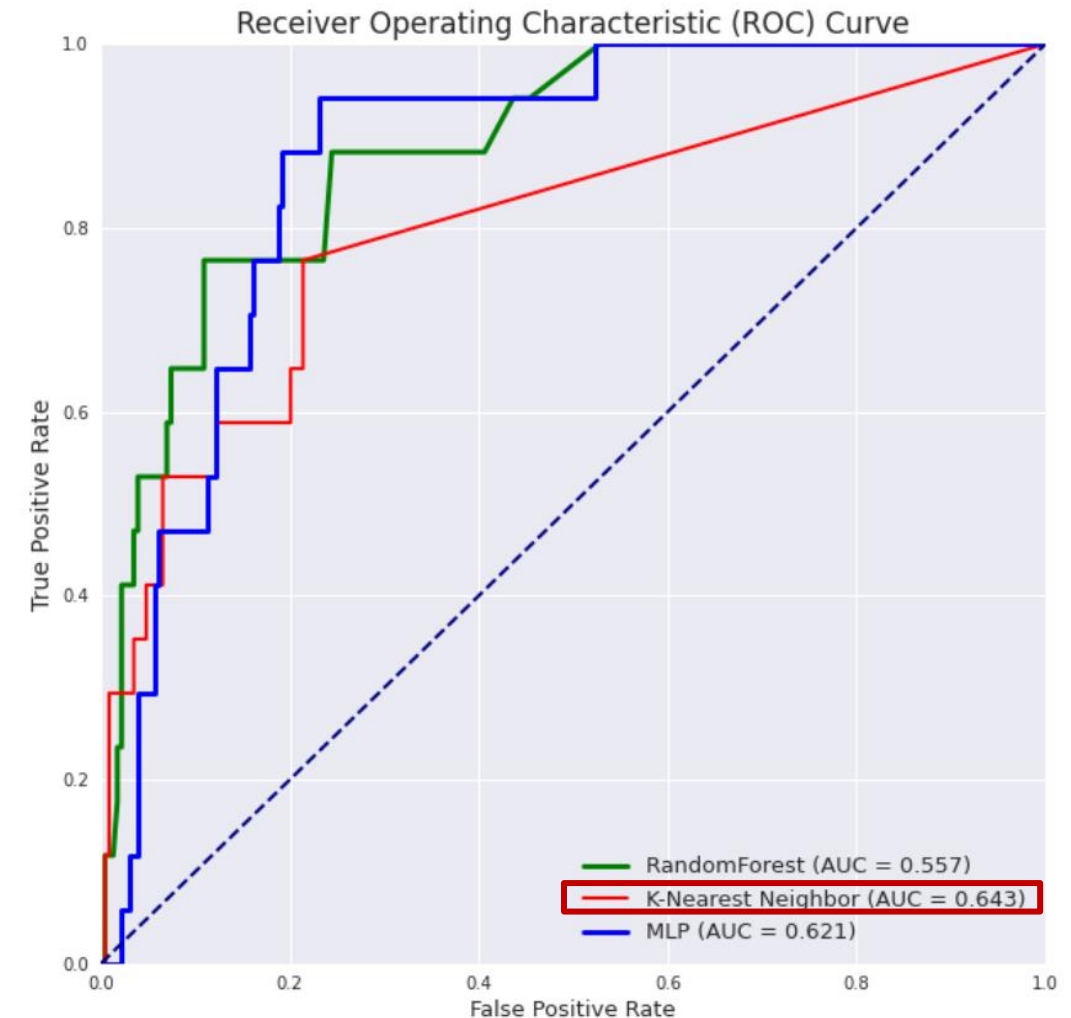
Book review statistics based on
the ISBNs list from **data.world**

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

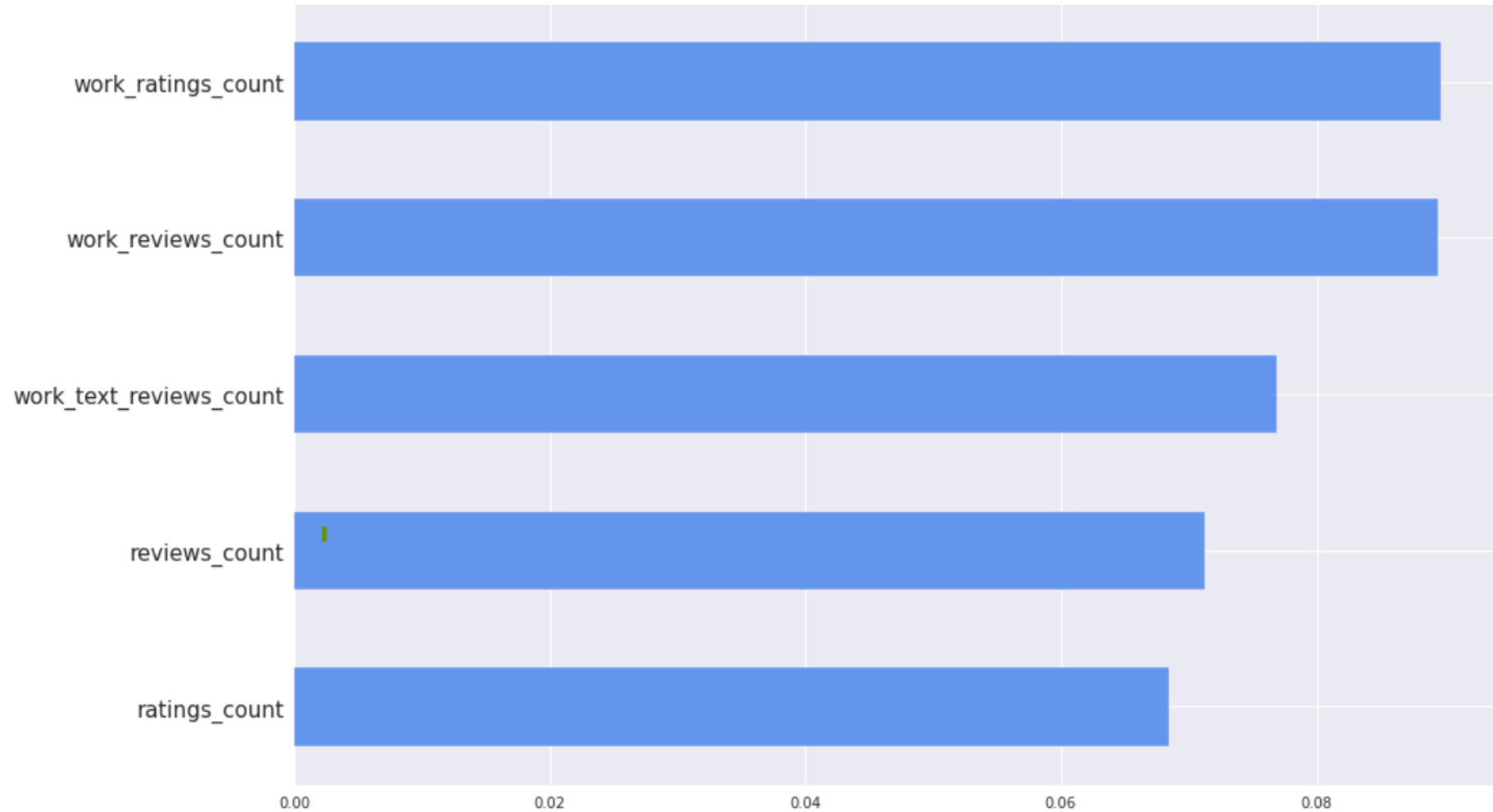


top 3 models' performance

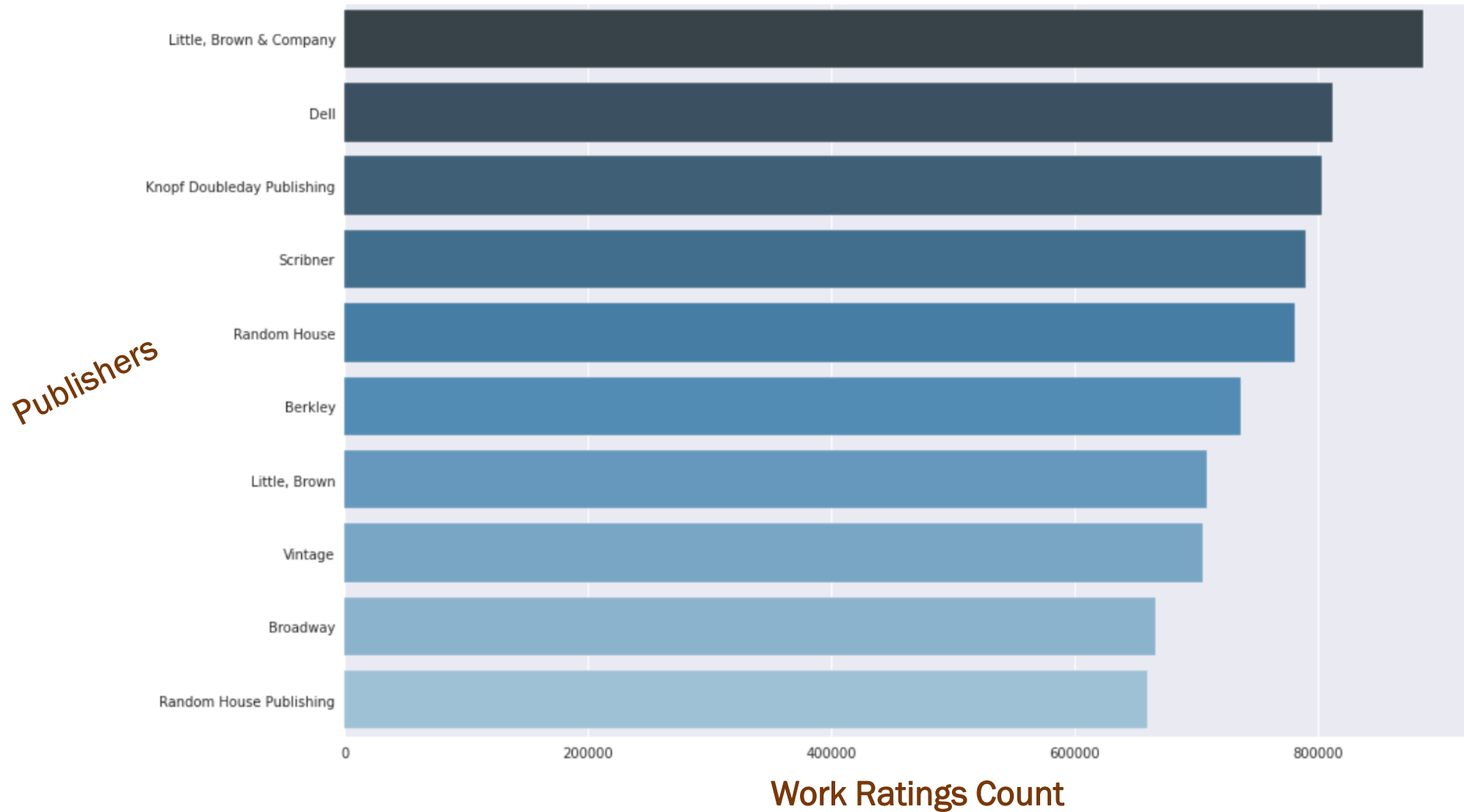
	Precision	AUC
<u>K-Nearest Neighbor</u>	0.71	0.64
<u>Random Forest</u>	0.67	0.56
<u>MLP</u>	0.28	0.62



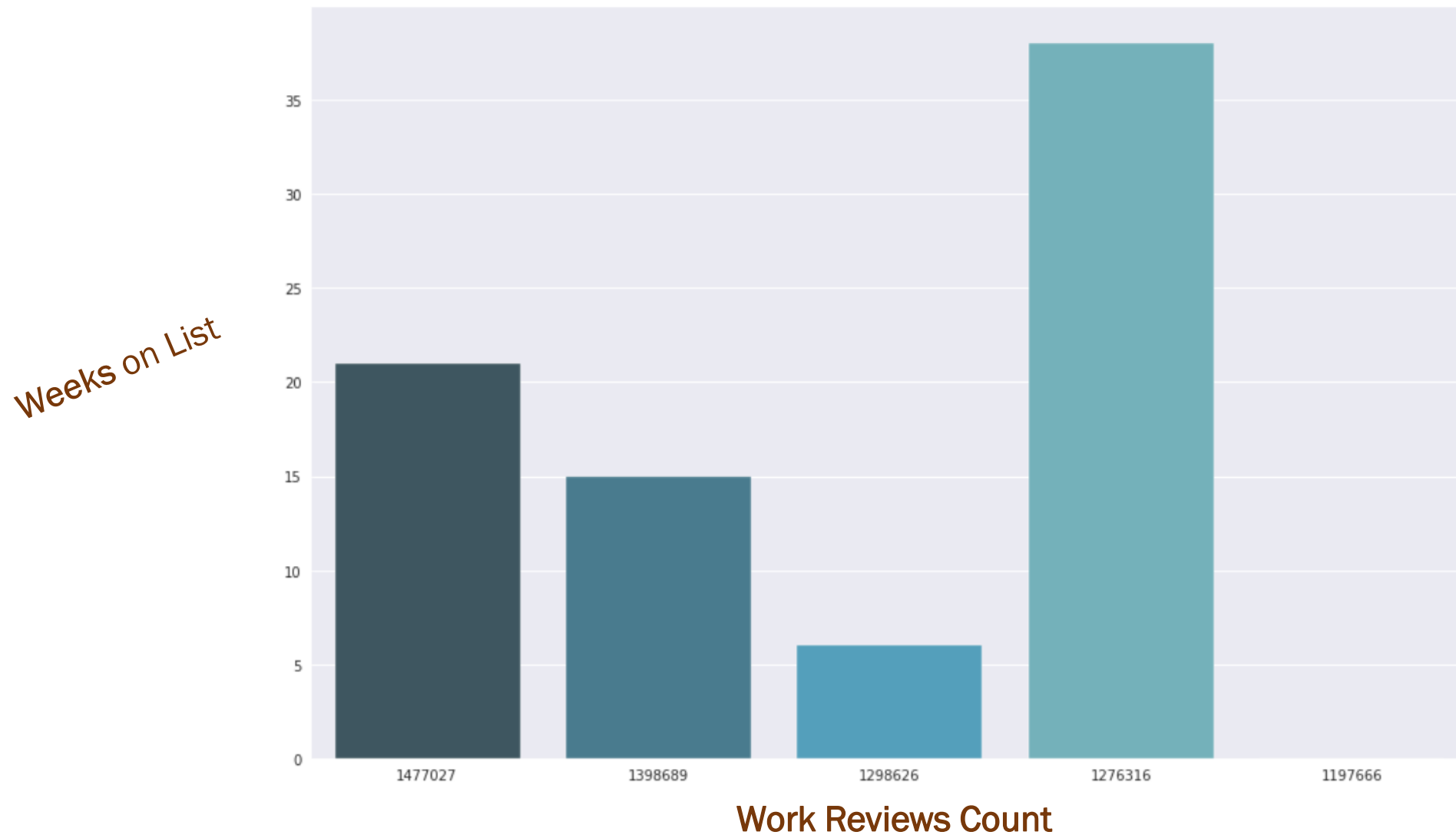
important features



recommendations



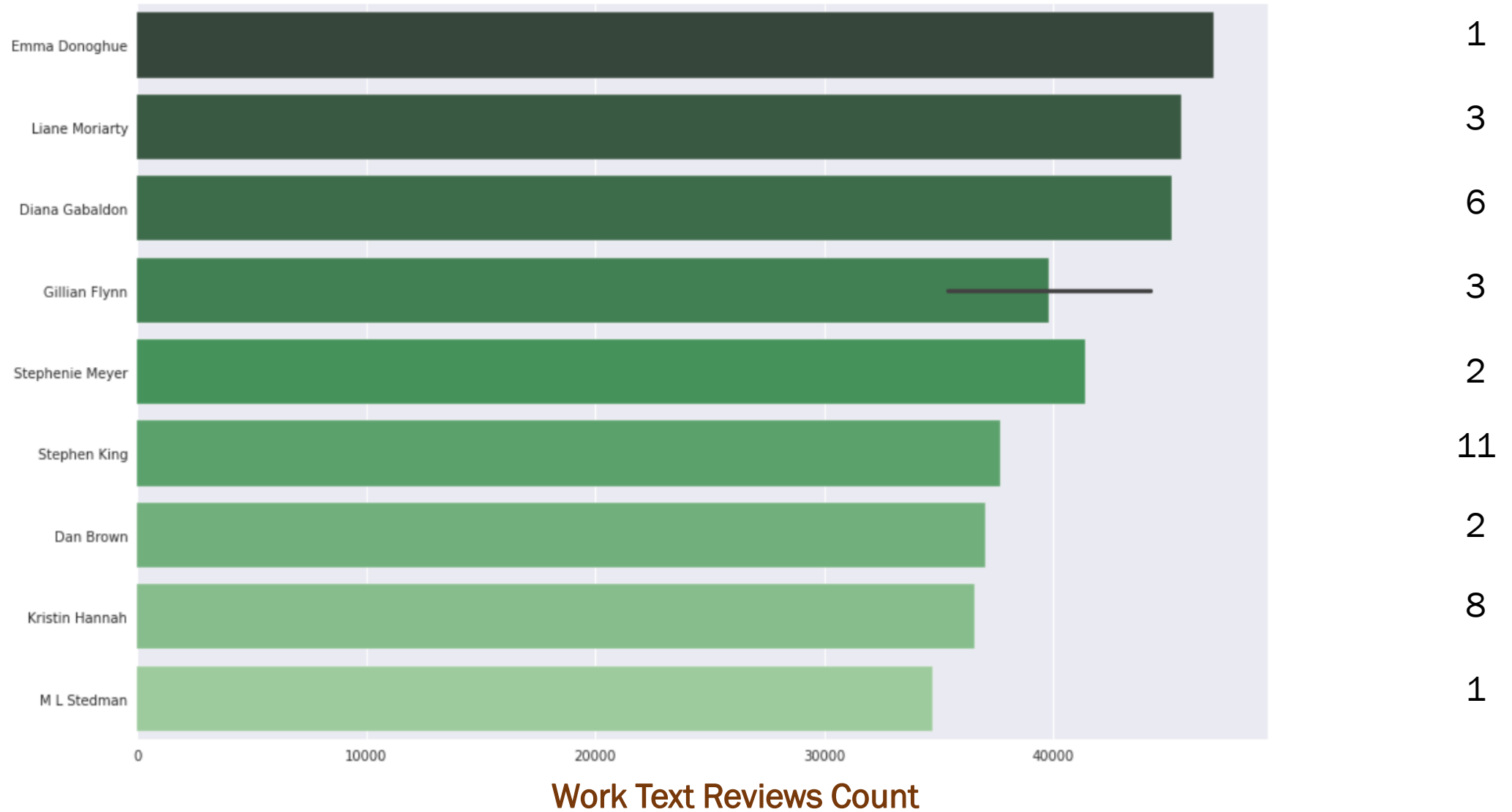
recommendations



recommendations

Number of books

Authors



recommendations

Work Reviews Count
&
Titles



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