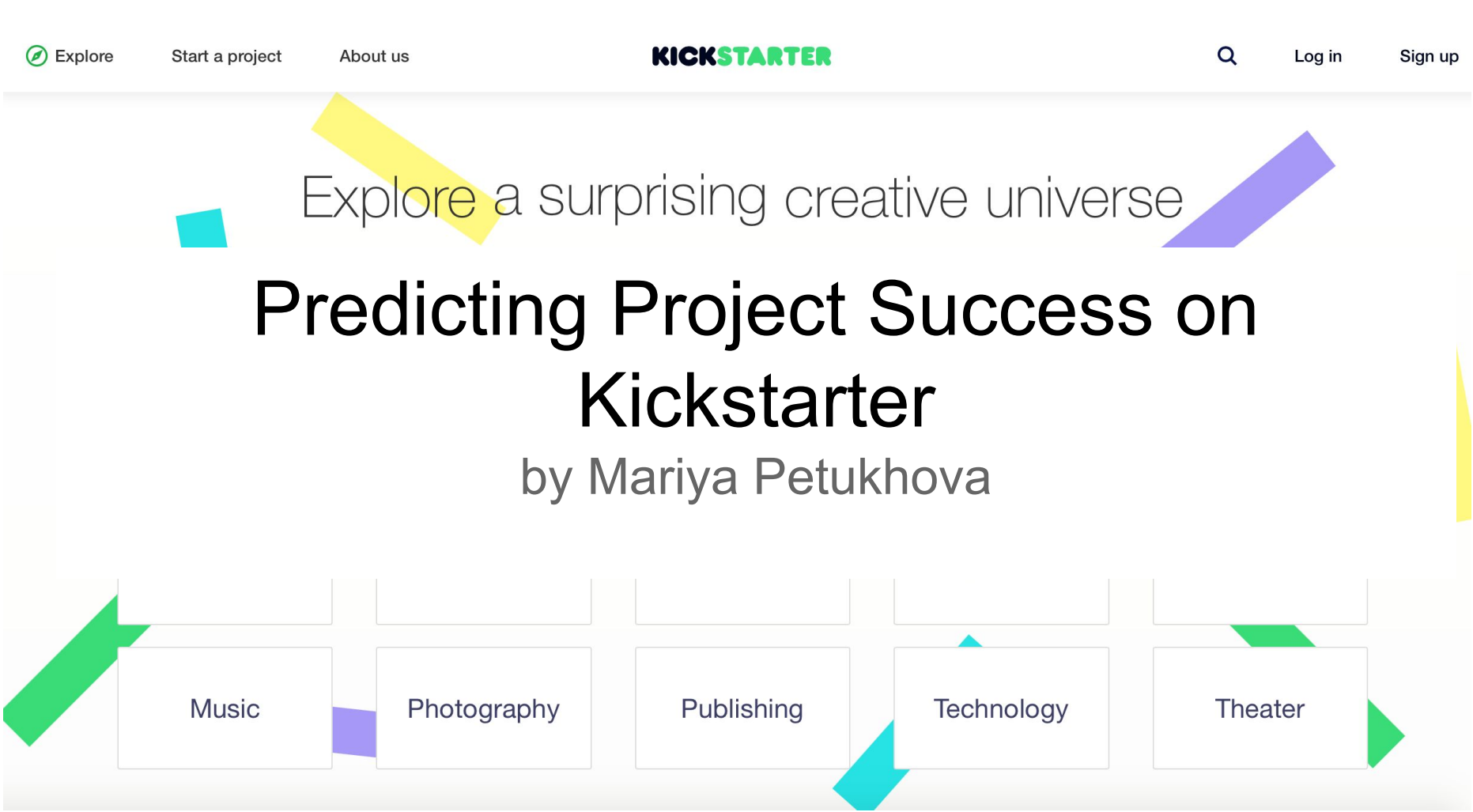

Explore
Start a project
About us
KICKSTARTER
Log in
Sign up
Explore a surprising creative universe
Predicting Project Success on Kickstarter
by Mariya Petukhova
Music
Photography
Publishing
Technology
Theater

Task

Can project’s pledged amount be predicted using other project features available through scraping of Kickstarter’s website?

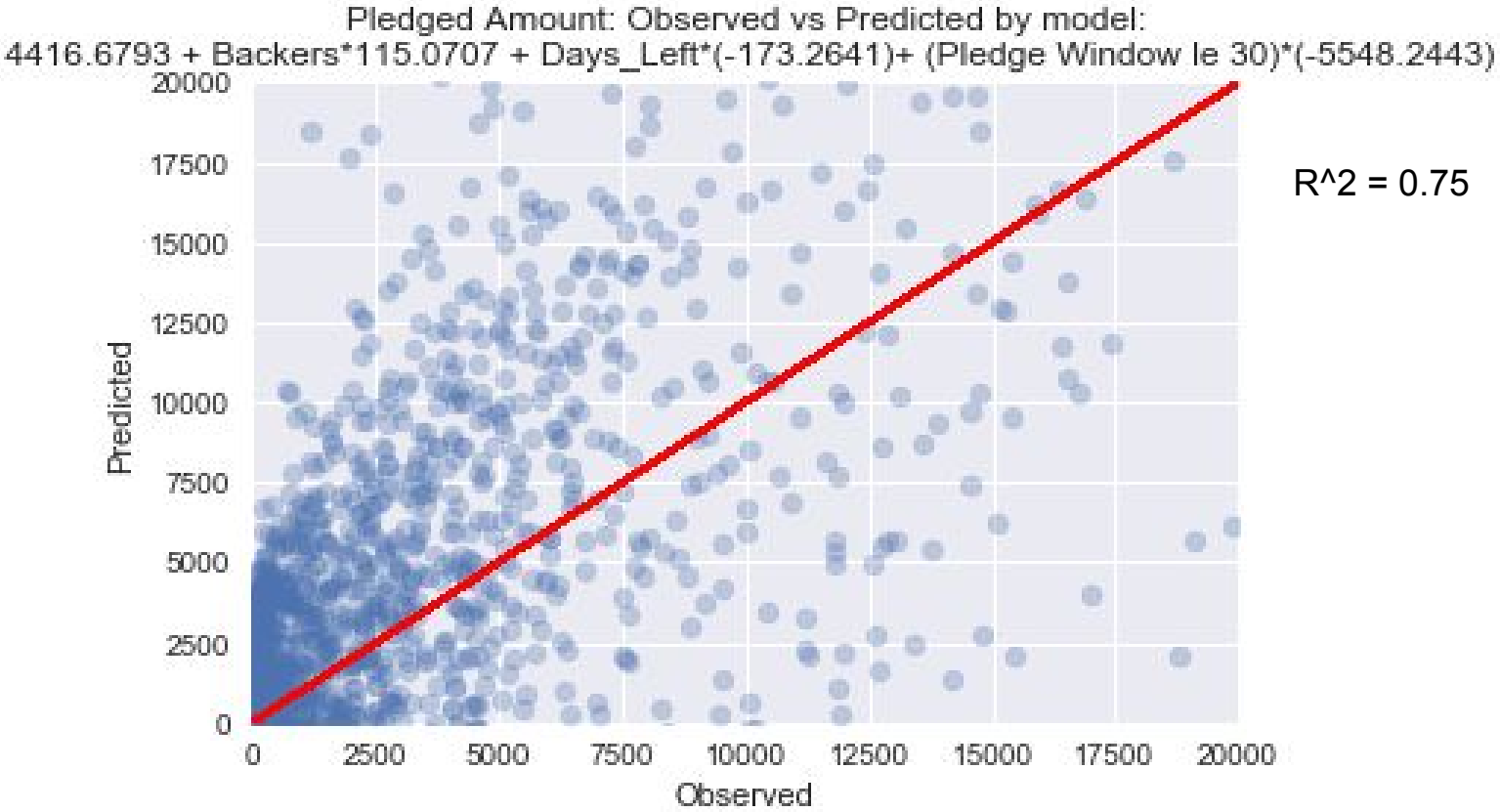
Data (web-scraped from Kickstarter.com on 3,169 projects)

- ❏ Goal Amount
- ❏ Pledged Amount
- ❏ Number of Backers
- ❏ Country/Location
- ❏ “Project We Love” Badge
- ❏ Project’s Launched and End Date

: “XML is a bitch to scrape.”

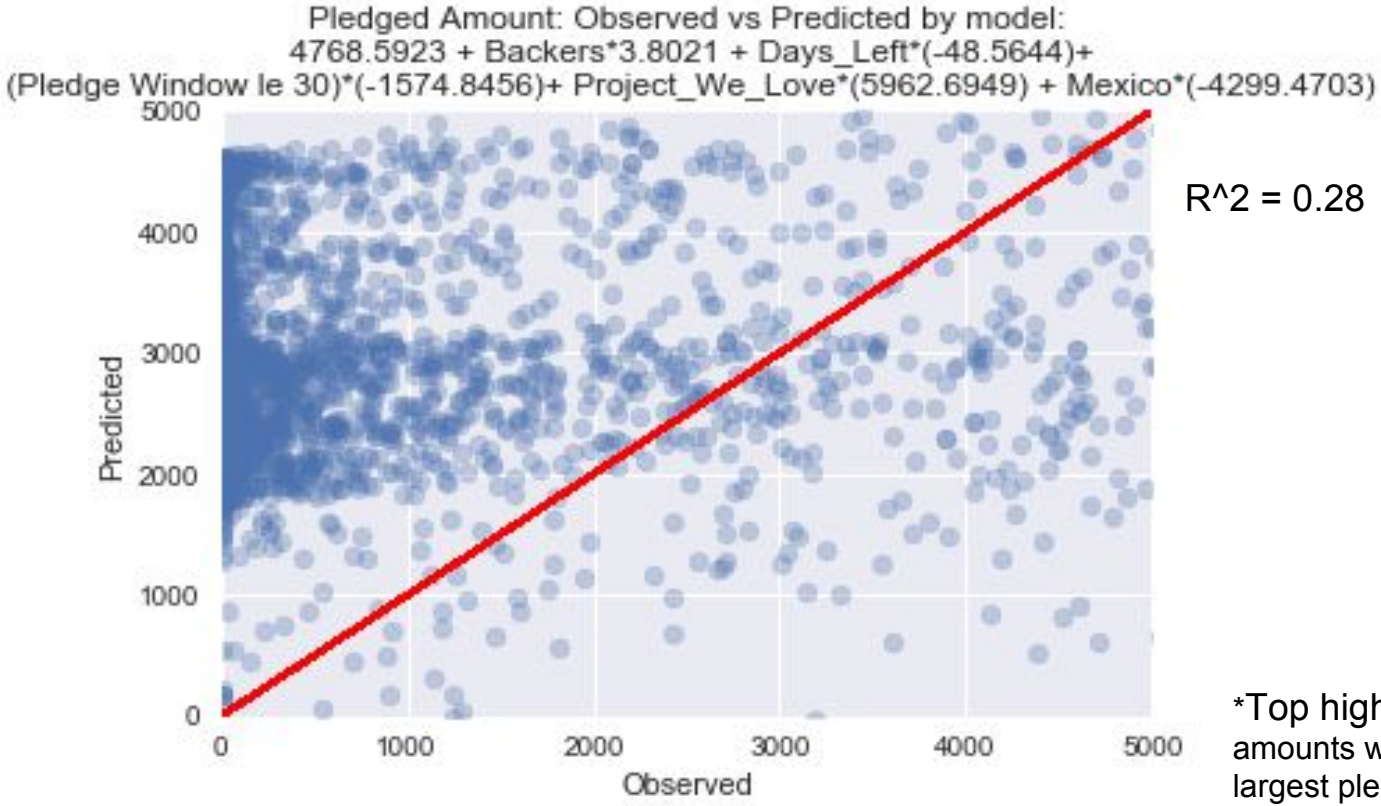
Word.

Model 1.0: Straight Up



R^2 = 0.75

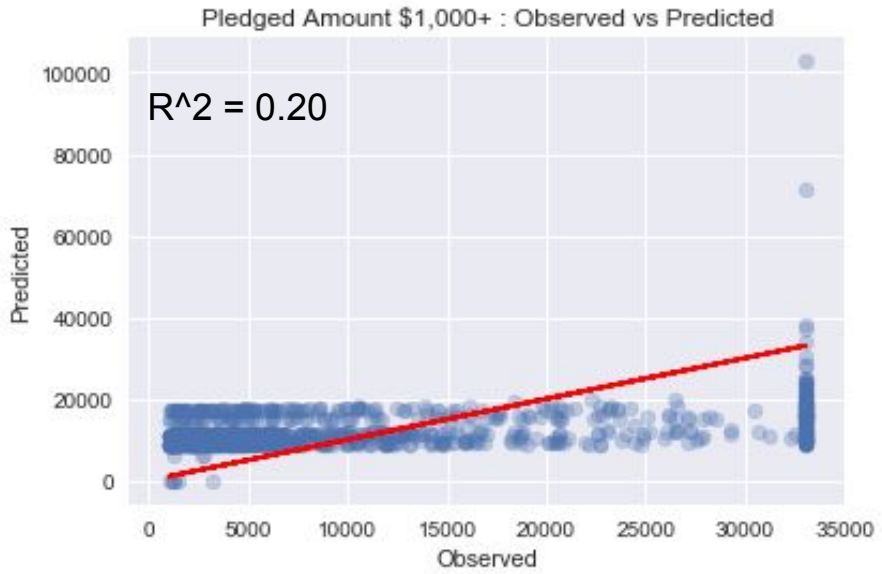
Model 2.0: Windsorized* Target

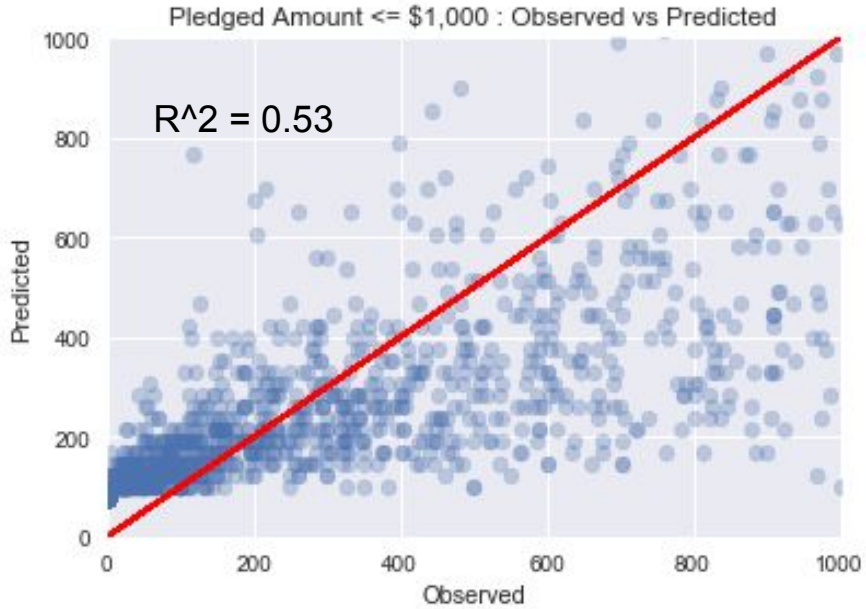


R^2 = 0.28

*Top highest 5% of pledged amounts was replaced with 95th largest pledge amount.

Models 3.0





- Pledged Amounts under $1,000 behave differently than the ones above $1,000 and two models capture it better
- A more complex methodology would suit this data better than Linear Regression

Conclusions

- Smaller Pledged Amounts are affected by number of backers only (from available features)
- Larger Pledged Amounts impacted by location, “Projects we love” badge, and whether the pledge time window was less than a month

Next steps:

- Collect larger sample
- Perform Text Analysis
- Collect data on funded projects performance post-Kickstarter
- Try supervised classification algorithm instead of Linear Regression


You can ask me anything you want,
ask me anything.

Appendix

Model: Pledged Amount Greater Than $1,000

	coef	std err	P>|t|
Intercept	8774.6897	421.183	0.000
Backers	3.1592	0.243	0.000
Project_We_Love	5962.6949	645.272	0.000
Mexico	-8883.2598	2884.732	0.002
Pledge Time Window at most 30 days	-2218.4727	546.601	0.000

Model: Pledged Amount No More Than $1,000

	coef	std err	P>|t|
Intercept	76.4606	5.331	0.000
Backers	22.9505	0.524	0.000