Dataset - csv consisting of approximately 273,000 kickstarter projects

Source - github repository: https://github.com/benrugg/kickscraper

Features of Interest: Kickstarter_id, Name, Slug, Blurb, Main category, Sub category, Launch date, Deadline date, Result, Goal, Pledged, Percent funded, Backer count, Currency

We would like to make a dashboard allowing the user to see how different features of kickstarter campaigns impact their probability of success. We will use Python, Pandas, Flask, a SQLite database, SQLAlchemy, HTML/CSS/Bootstrap, Javascript, and Plotly.js to build this dashboard.

Potential questions that could be answered by exploring this dashboard:
What categories are most successful?
What subcategories are most successful within a given category?
What is the ideal length of time for a kickstarter project?
Has the success rates of kickstarter campaigns changed over time (how does the launch date impact success)?
How large can you make your goal and still be successful?
How many backers do you need to attract to be successful?

As a stretch goal we will see if we can train a logistic machine learning model to give the probabilities of the possible outcomes for a kickstarter campaign based on the following variables: category, subcategory, campaign length, goal amount, currency, blurb length, and blurb sentiment values. This could let a user input a proposed kickstarter campaign idea to predict the probability that the campaign will be successful.

We could potentially have a model with backer count and without backer count. This seems like the hardest thing to predict before you start your campaign. If you had a strategy in place to attract a certain number of people this could be a very helpful prediction metric though.

We could also potentially see if we could use clustering machine learning algorithm to see if we could predict the category and subcategory of a kickstarter campaign based on the blurb. This could let a user check to see if they are going to submit their campaign within the right category.