UC Berkeley Extension – Data Analytics Boot Camp

Module 1

Q1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

* Most created category (“Theater”) vs least created category (“journalism”).
* Most created Sub-category (“Play”) vs least created sub-category (“world music”).
* This is pretty obvious but, more backers you have, better chance of achieving the goal.
* Even if the backer number is not the highest, if number of average donations are high, better chance of achieving the set goal.
* Q3 (July to Sept) time of the year has one of the highest numbers of decline/drop in both success and fail numbers.
* “Journalism” category has 100% percentage of achieving the goal.
* “world music” and “audio” sub-category also has 100% percentage of achieving the goal.

Q2. What are some limitations of this dataset?

* Limited Number of country of data (only 7 countries)
* Limited Variety of data from each year (There are only 2 data from 2020 year for launched date).
* Lack of information on backers (age, gender, etc)

Q3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

* Percentage of failed vs successful rate by each category or sub-category.
* Date/Time line of created to ended (how many days/week/months)