**Individual Challenge 1: Excel**

Ricardo De los Rios

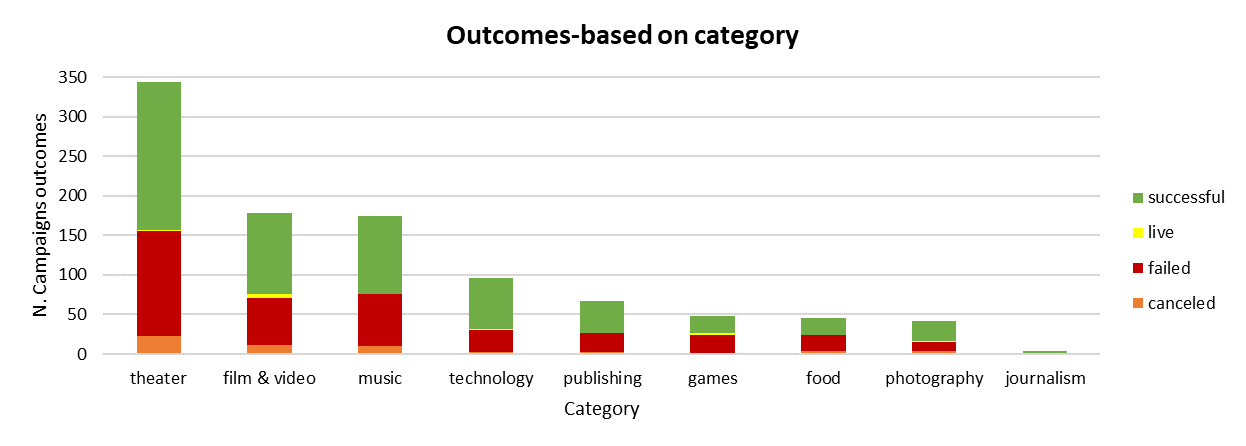
Data Analytics Boot Camp

June, 2023

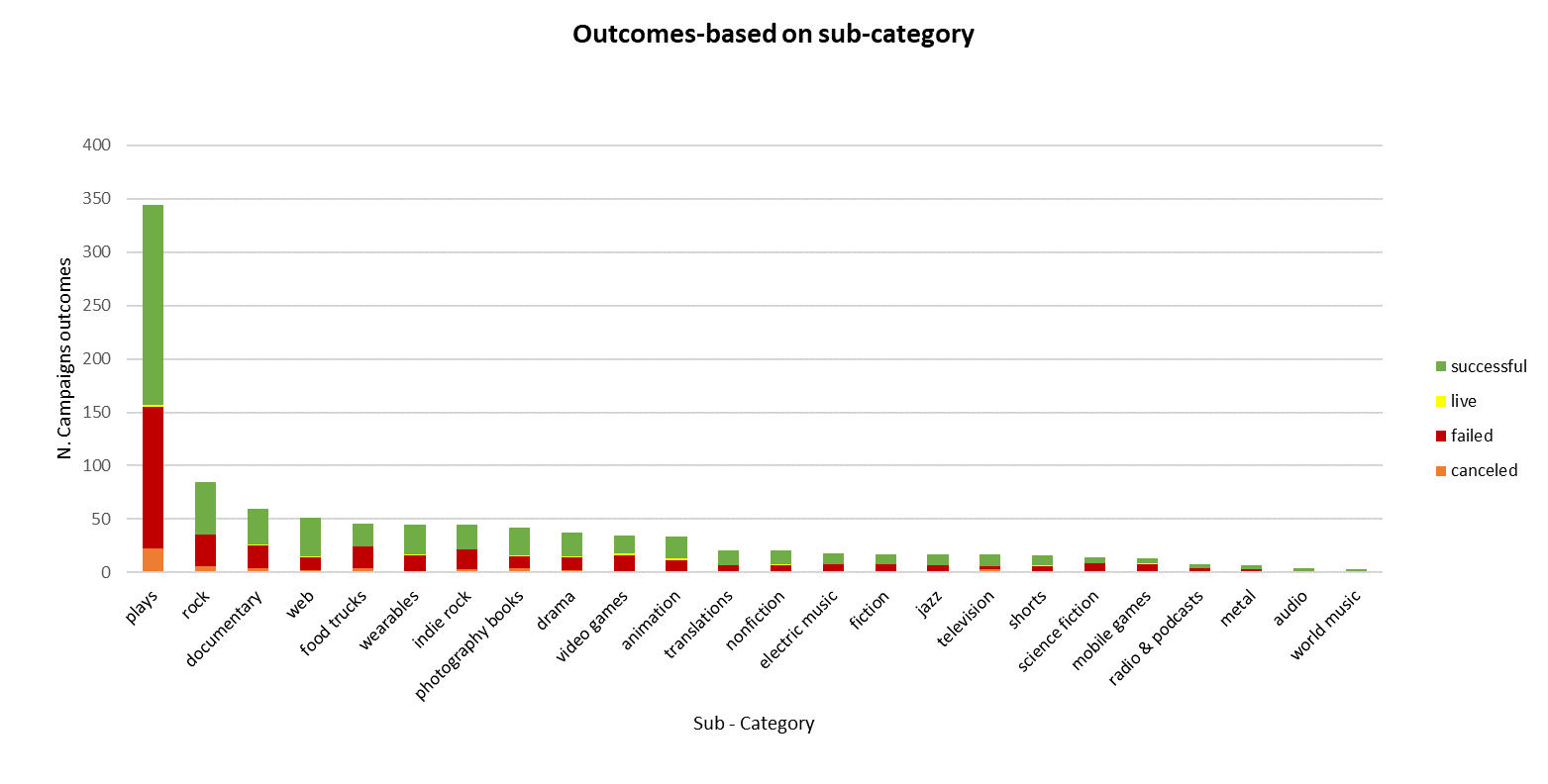
**Module 1 Challenge**

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

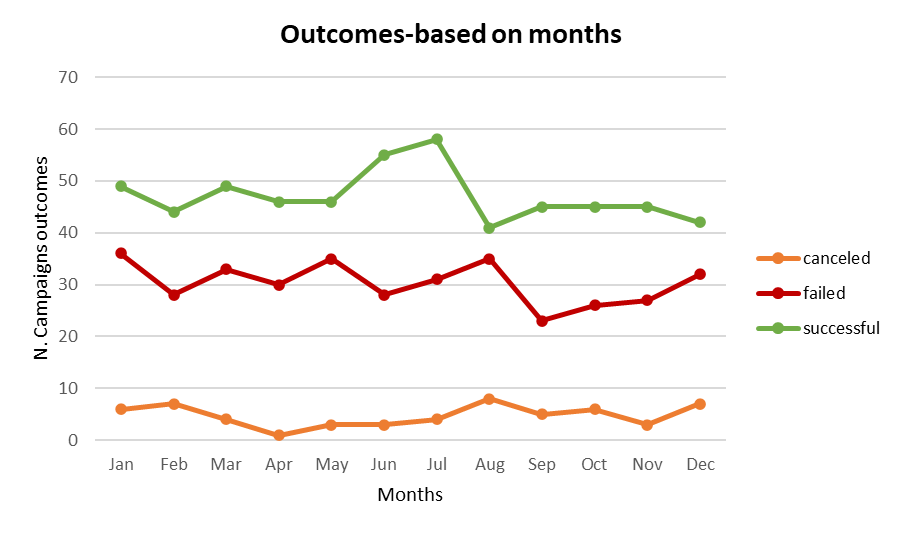
According to the information from the crowdfunding campaigns carried out, it can be concluded that of the total of 565 successful campaigns, 80% of the campaigns were carried out in 4 categories: theater 187 (33.09%), video and film 102 (18.05%), music 99 (17.52%) and technology 64 (11.32%). Likewise, it observes that in the same categories, the percentage of success was higher than 50%, being theater (54%), video and film (57%), music (57%), and technology (67%).



Regarding the analysis by subcategory, it can be concluded that plays with 344 campaigns represents 34% of the total subcategories; likewise, it can be seen that of the total number of successful campaigns (565), 187 is 33%, corresponding to the plays subcategory. In terms of failed campaigns, the plays sub-category was the one that had the most with 132 out of 364, which represents 36%, and in terms of the highest number of canceled campaigns, it was also the subcategory called plays that registered the highest number with 23 out of 57, which is equivalent to the 40%.



In the trend that appears in the following graph, the month with the highest number of successful outcomes was July with 58. On the other hand, August was the month with the highest number of failed outcomes with 35, and canceled with 8.



* + **What are some limitations of this dataset?**

One of the limitations that is evident in the database is that there are variables that are not standardized, such as currency. This is found in different types such as USD, CAD, AUD, CHF, DKK, EUR, GBR; which limits the analysis in monetary terms.

Additionally, it can say that the design of the database makes the analysis more complex; For instance, if that database will reach a person who does not know how to convert the variables (launched\_at and Deadline) from timestamp to data, they will surely lose a lot of information when making a report. Another case of complexity that it can observe is the way in which the information of (category & sub-category) is displayed since like the previous example, it requires additional knowledge to separate the information into categories and sub-categories. The databases should be easy to understand and facilitate the work of any user who requires the information.

Finally, another limitation that people would find is that by having to insert columns or variables to perform additional calculations, the integrity and precision of the information may be affected since by allowing manipulation and/or modification, human errors or transfer errors may appear.

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

First, a table and graph could be created showing how many campaigns were successful, failed, or cancelled, or are currently active by country. Because with this information it is possible to identify which is the country that carries out the most campaigns.

Second, one could go deeper by country with another table and graph to identify which categories and subcategories are chosen in the country that carries out the most campaigns.

Third, it could include a dynamic table of the Top 5 of the most successful categories with their respective subcategories and thus determine which strategy is applied to the category.

Finally, if the currency was standardized, an analysis of which category and sub-category more resources were received, would allow directing future strategies to get more resources.