COLUMBIA UNIVERSITY

DATA ANALYTICS BOOTCAMP

NEW YORK

Crowdfunding Goal Analysis

Statistical Analysis

Profesor: Jean-Paul Lavalley

IVAN GALEANO

December 21, 2022

Table of Contents

[Crowdfunding Goal Analysis: 3](#_Toc526602617)

Three conclusions from the data[. 4](#_Toc526602619)

limitations of the dataset and suggestions for additional tables of graph [. 5](#_Toc526602620)

A brief and compelling justification of whether the mean or median better summarizes the data [6](#_Toc526602621)

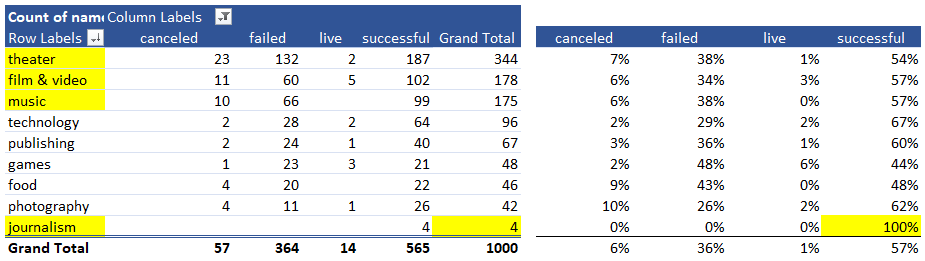
# Crowdfunding Goal Analysis:

After a exhaustive data analysis there are few conclusions that can be inferred from the data provided:

* The US is the leader in crowdfunding initiative for projects with 76%.
* The three major categories for crowdfunding projects are related to arts (Theater, Film and video, Music) and this is true for the US and for all the other countries in the dataset.
* Journalism projects were 100% successful although these types of projects only represent a 4% of the total projects

A picture containing waterfall chart

Description automatically generated



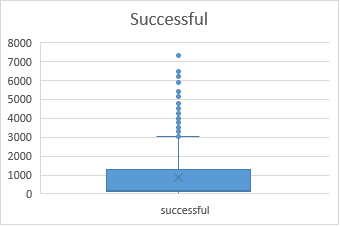
# limitations of the dataset and suggestions for additional tables of graph

The crowdfunding data analyzed seem to be from only one source and the data set rather small given that we find data points from different countries, and the spam of time is rather large for only 1000 records.

Another characteristic of the data is that in order to arrive to meaningful conclusion related to the amounts of money involved, we should have an exchange rata for the other currencies that are not US dollars.

Other elements that can be graph is the relationship of the most successful categories through the years to identify if the success was due to some new trend at one point in time or if in fact these are categories that are most likely to be successful.

# Compelling justification of whether the mean or median better summarizes the data



The best way to understand our data is to use a Box and whiskers graph, because we will be able to see what data points are outliers that only bring distraction to what our data really represents. In the graph we can see that our Mean is at the Upper Quartile of the graph, which will make the Median the one that better summarizes the data.