VIK SHAH

14th Feb 2022

Excel Assignment Charting Crowdfunding

A screenshot of a video game

Description automatically generated with medium confidence

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

* Theatre/ Plays have the most crowdfunding campaigns
* Journalism have the least crowdfunding campaigns
* There are more successful campaigns than failed or canceled
* More people crowdfund in US than any other country from our dataset

2. What are some limitations of this dataset?

* Multiple currencies, with no exchange rate hence amounts may not be comparable between countries
* What other influences existed to contribute to the crowdfunding donations e.g. advertising, promotions, needs based application of campaign that influence pledges
* Are we to assume these are all original crowd funding campaigns? Or some may exist on our data sources or have previously been cancelled and setup again under another stream, hence the proportion of cancelled/ failed campaigns to successful ones may be misleading. E.g. if campaign x wants to raise $100 and has raised $60 on one platform and $40 on another, the data from a single platform may misguide the reader
* Number of users on crowdfunding platform in 2010 vs 2020. In 2010 successful crowdfunding would’ve received a different target audience/ demographic than 2020 which may aid raising awareness. As technology became more widespread and smartphones became everyone’s essential, during the smart phone era, understanding and use of crowdfunding may have evolved.

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

* Number of days to achieve a successful campaign – may be useful for optimising time/ resource to host/ promote crowdfunding campaign
* Average Donation by category – may be useful as a helpful funding target request
* Country Ratio & contribution – may be useful to determine resource allocation and time spent in promoting campaigns in other countries

**Bonus**

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.

* Mean and median without mode aren’t all the measures of central tendency and thus a bit limiting. These are meaningful simply as reference points to interpret that large dataset.
* The std deviation is greater than mean so the data points are further from the mean, hence a higher deviation within the data set. i.e. the more spread out the data, more variance to take into considerations in conclusions
* Introducing box & whisker plots does assist in determining outliers.

2. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

There is more variance with successful campaigns, given it has a larger standard deviation (from the formulas). Successful campaign has more outliers than failed campaign (found from counting campaigns outside inter quartile range). This would contribute to the variability.