**Background**

Crowdfunding platforms like Kickstarter and Indiegogo have been growing in success and popularity since the late 2000s. From independent content creators to famous celebrities, more and more people are using crowdfunding to launch new products and generate buzz, but not every project has found success.

To receive funding, the project must meet or exceed an initial goal, so many organizations dedicate considerable resources looking through old projects in an attempt to discover “the trick” to finding success. For this week's Challenge, you will organize and analyze a database of 1,000 sample projects to uncover any hidden trends.

**Q&A**

* Create a report in Microsoft Word, and answer the following questions:
  + Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
* Number of successfully approved projects peaks at June and July
* Number of failed projects peaks at August. It follows low approval rate for successful projects as well.
* Number of cancelled projects is relatively low every month however if peaks around August as well.
  + What are some limitations of this dataset?
* We only looked at 10 years’ worth of data
* We have relatively small dataset per country. The approval rate could depend on the country.
* We should look at the age of applicants. Are younger applicants more likely to get approved for funding than an older applicates?
* Do gender and race have an impact?
* It would be nice to have a summary of why the project was successful, failed or canceled to see if there are any trends.
  + What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
* We can create a stacked pivot table to see if the approval trend is the same every year
* We can look at the percent of successful, failed and canceled projects to normalize the data.
* We can look at the approved funding per the amount of funding specified in the Goal

**Statistical Analysis Q&A**

* Use your data to determine whether the mean or the median better summarizes the data.
* For this data set mean is significantly different from a median for both successful and failed outcome. Median better summarizes this data because it tells me that for 50% of the projected need less or equal number of backers and for 50% of the projected need more or equal number of backers.
* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
* As expected, the variability in data is higher for successful projects than for failed projects because both variance and standard deviation are higher for successful projects. This means larger number of backers was willing to invest into some of the successful projects and lower number of backers was willing to invest into some of the failed projects.

|  |  |  |
| --- | --- | --- |
| **Statistical Analysis** | **outcome\_successful** | **outcome\_failed** |
| The mean number of backers | 851 | 586 |
| The median number of backers | 201 | 115 |
| The minimum number of backers | 16 | 0 |
| The maximum number of backers | 7295 | 6080 |
| The variance of the number of backers | 1606217 | 924113 |
| The standard deviation of the number of backers | 1267 | 961 |

**Requirements**

**Conditional Formatting (10 points)**

* Conditional formatting is applied appropriately to the outcome column (5 points)
* Conditional formatting is applied appropriately to the percent funded column (5 points)

**Column Creation (10 points)**

* Six new columns were correctly created for:
  + percent funded
  + average donation
  + category
  + sub-category
  + Date Created Conversion
  + Date Ended Conversion

**Pivot Tables and Stacked Column Charts (15 points)**

* Correctly created a pivot table that counts how many campaigns were "successful," "failed," "canceled," or are currently "live" per category (7.5 points)
* Correctly created a stacked column pivot chart that can be filtered by country (7.5 points)

**Pivot Tables and Line Graphs (15 points)**

* Correctly created a pivot table with a column of outcome, rows of Date Created Conversion, values based on the count of outcome, and filters based on parent category and Years (7.5 points)
* Correctly created a pivot chart line graph (7.5 points)

**Written Report (20 points)**

* Presents a cohesive written analysis that:
  + Draws three conclusions from the data (10 points)
  + States limitations of the dataset and suggestions for additional tables of graph (10 points)

**Crowfunding Goal Analysis (10 points)**

* Computed calculations of percentages for projects that were successful, failed, or were canceled per goal range (5 points)
* Created a line chart showing the relationship between the goal’s amount and its chances at success, failure, or cancellation (5 points)

**Statistical Analysis (20 points)**

* Computed calculations of the mean, median, min, max, variance, and stdev using Excel formulas (15 points)
* A brief and compelling justification of whether the mean or median better summarizes the data (5 points)

**Grading**

This assignment will be evaluated against the requirements and assigned a grade according to the following table:

| **Grade** | **Points** |
| --- | --- |
| A (+/-) | 90+ |
| B (+/-) | 80–89 |
| C (+/-) | 70–79 |
| D (+/-) | 60–69 |
| F (+/-) | < 60 |

**Submission**

To submit your Challenge assignment, click Submit, and then provide the URL of your GitHub repository for grading.

**NOTE**

You are allowed to miss up to two Challenge assignments and still earn your certificate. If you complete all Challenge assignments, your lowest two grades will be dropped. If you wish to skip this assignment, click Next, and move on to the next Module.

Comments are disabled for graded submissions in BootCamp Spot. If you have questions about your feedback, please notify your instructional staff or your Student Success Manager. If you would like to resubmit your work for an additional review, you can use the Resubmit Assignment button to upload new links. You may resubmit up to three times for a total of four submissions.

**IMPORTANT**

No matter how difficult the course becomes, you must always turn in original work. Plagiarism is not tolerated. If your instructional or support staff determine that you have plagiarized work, your Student Success Manager will determine the appropriate course of action based on university policy. Such actions may include, but are not limited to, a documented plagiarism discussion, an incomplete or failing grade assignment, or ineligibility for graduation.

**It is your responsibility to include a note in the README section of your repo specifying code source and its location within your repo**. This applies if you have worked with a peer on an assignment, used code in which you did not author or create sourced from a forum such as Stack Overflow, or you received code outside curriculum content from support staff such as an Instructor, TA, Tutor, or Learning Assistant. This will provide visibility to grading staff of your circumstance in order to avoid flagging your work as plagiarized.

**References**

Data for this dataset was generated by edX Boot Camps LLC, and is intended for educational purposes only.