Module 1 Challenge

Start Assignment

Due Jun 15 by 11:59pm **Points** 100 **Submitting** a text entry box or a website url

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.

Before You Begin

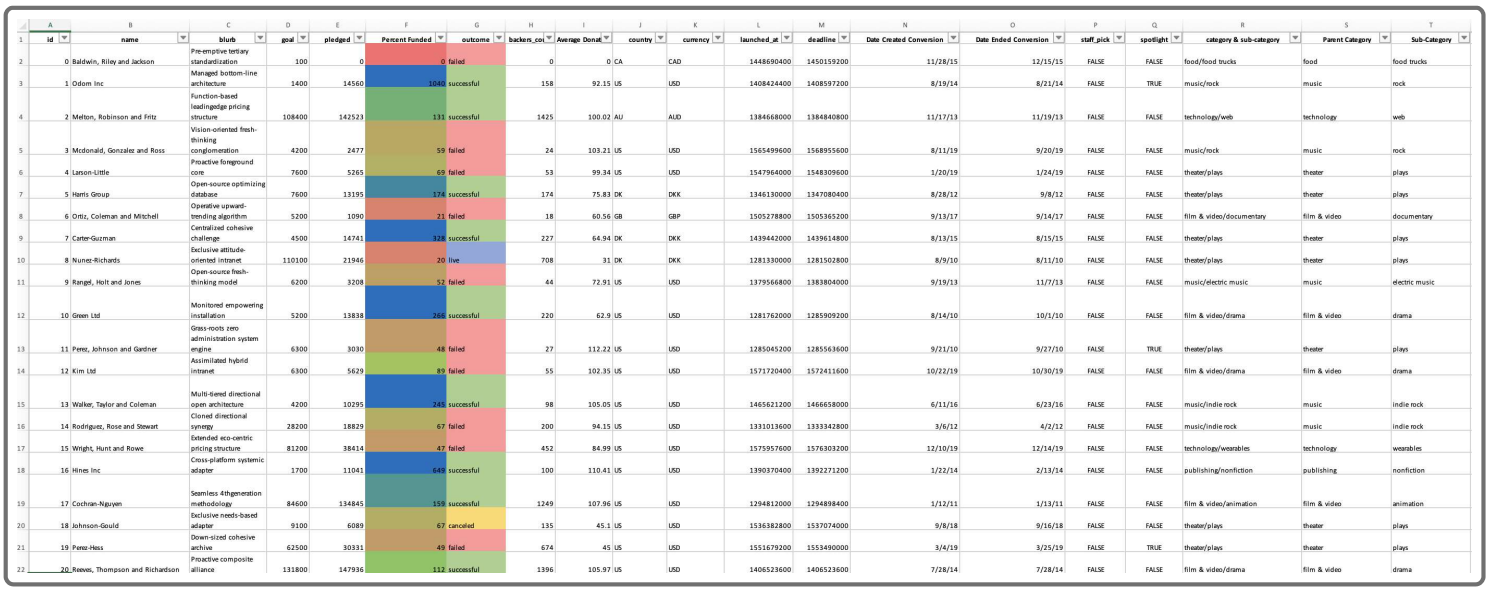
1. Create a new space for this project called `excel-challenge` in either Dropbox or Google Drive. **Do not add this Challenge to an existing repository**.
2. Store your Excel workbooks here in this new space, and create a sharable link for submission.

Files

Download the following files to help you get started:

Module 1 Challenge files (https://static.bc-edx.com/data/dl-1-2/m1/lms/starter/Starter_Code.zip)

Instructions

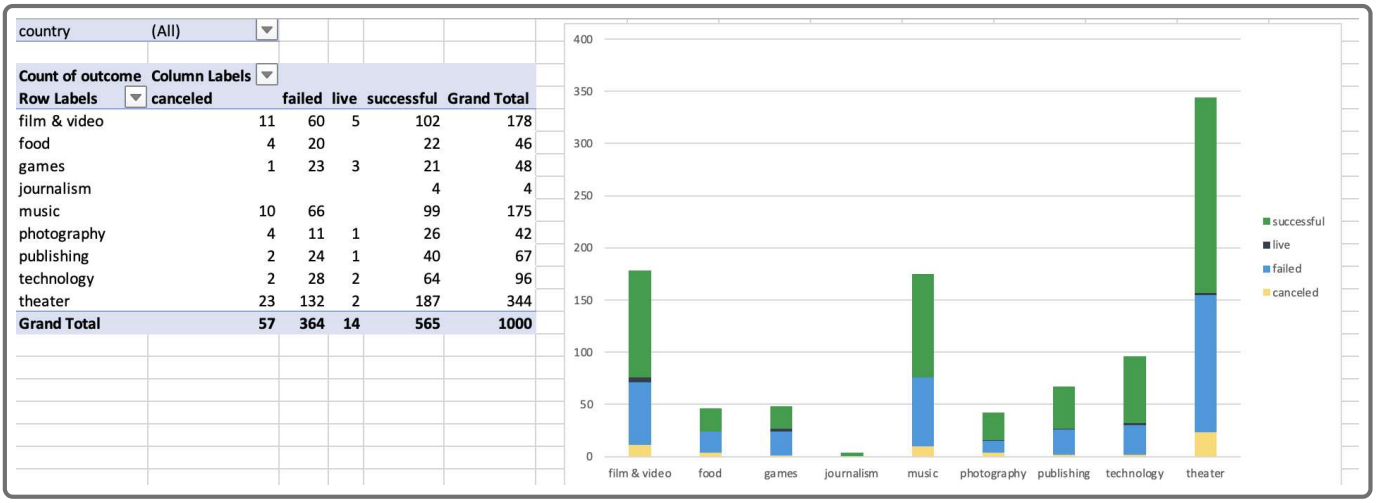


id	name	blurb	goal	pledged	Percent Funded	outcome	backers_co	Average Donat	country	currency	launched_at	deadline	Date Created Conversion	Date Ended Conversion	staff_pick	spotlight	category & sub-category	Parent Category	Sub-Category
0	Baldwin, Riley and Jackson	Pre-emptive tertiary standardization	100	0	0	failed	0	0	CA	CAD	1448690400	1450159200	11/28/15	12/15/15	FALSE	FALSE	food/food trucks	food	food trucks
1	Odom Inc	Managed bottom-line architecture	1400	14560	1040	successful	158	92.15	US	USD	1408424400	1408597200	8/19/14	8/21/14	FALSE	TRUE	music/rock	music	rock
2	Melton, Robinson and Fritz	Function-based leadingedge pricing structure	108400	142523	131	successful	1425	100.02	AU	AUD	1384668000	1384840800	11/17/13	11/19/13	FALSE	FALSE	technology/web	technology	web
3	Mcdonald, Gonzalez and Ross	Vision-oriented fresh-thinking conglomeration	4200	2477	59	failed	24	103.21	US	USD	1565499600	1568955600	8/11/19	9/20/19	FALSE	FALSE	music/rock	music	rock
4	Larson-Little	Proactive foreground core	7600	5265	69	failed	53	99.34	US	USD	1547964000	1548309600	1/20/19	1/24/19	FALSE	FALSE	theater/plays	theater	plays
5	Harris Group	Open-source optimizing database	7600	13195	174	successful	174	75.83	DK	DKK	1346130000	1347080400	8/28/12	9/8/12	FALSE	FALSE	theater/plays	theater	plays
6	Ortiz, Coleman and Mitchell	Operative upward-trending algorithm	5200	1090	21	failed	18	60.56	GB	GBP	1505278800	1505365200	9/13/17	9/14/17	FALSE	FALSE	film & video/documentary	film & video	documentary
7	Carter-Guzman	Centralized cohesive challenge	4500	14741	328	successful	227	64.94	DK	DKK	1439442000	1439614800	8/13/15	8/15/15	FALSE	FALSE	theater/plays	theater	plays
8	Nunez-Richards	Exclusive attitude-oriented intranet	110100	21946	20	live	708	31	DK	DKK	1281330000	1281502800	8/9/10	8/11/10	FALSE	FALSE	theater/plays	theater	plays
9	Rangel, Holt and Jones	Open-source fresh-thinking model	6200	3208	52	failed	44	72.91	US	USD	1379566800	1383804000	9/19/13	11/7/13	FALSE	FALSE	music/electric music	music	electric music
10	Green Ltd	Monitored empowering installation	5200	13838	266	successful	220	62.9	US	USD	1281762000	1285909200	8/14/10	10/1/10	FALSE	FALSE	film & video/drama	film & video	drama
11	Perez, Johnson and Gardner	Grass-roots zero administration system engine	6300	3030	48	failed	27	112.22	US	USD	1285045200	1285563600	9/21/10	9/27/10	FALSE	TRUE	theater/plays	theater	plays
12	Kim Ltd	Assimilated hybrid intranet	6300	5629	89	failed	55	102.35	US	USD	1571720400	1572411600	10/22/19	10/30/19	FALSE	FALSE	film & video/drama	film & video	drama
13	Walker, Taylor and Coleman	Multi-tiered directional open architecture	4200	10295	245	successful	98	105.05	US	USD	1465621200	1466658000	6/11/16	6/23/16	FALSE	FALSE	music/indie rock	music	indie rock
14	Rodriguez, Rose and Stewart	Cloned directional synergy	28200	18829	67	failed	200	94.15	US	USD	1331013600	1333342800	3/6/12	4/2/12	FALSE	FALSE	music/indie rock	music	indie rock
15	Wright, Hunt and Rowe	Extended eco-centric pricing structure	81200	38414	47	failed	452	84.99	US	USD	1575957600	1576303200	12/10/19	12/14/19	FALSE	FALSE	technology/wearables	technology	wearables
16	Hines Inc	Cross-platform systemic adapter	1700	11041	649	successful	100	110.41	US	USD	1390370400	1392271200	1/22/14	2/13/14	FALSE	FALSE	publishing/nonfiction	publishing	nonfiction
17	Cochran-Nguyen	Seamless 4thgeneration methodology	84600	134845	159	successful	1249	107.96	US	USD	1294812000	1294898400	1/12/11	1/13/11	FALSE	FALSE	film & video/animation	film & video	animation
18	Johnson-Gould	Exclusive needs-based adapter	9100	6089	67	canceled	135	45.1	US	USD	1536382800	1537074000	9/8/18	9/16/18	FALSE	FALSE	theater/plays	theater	plays
19	Perez-Hess	Down-sized cohesive archive	62500	30331	49	failed	674	45	US	USD	1551679200	1553490000	3/4/19	3/25/19	FALSE	TRUE	theater/plays	theater	plays
20	Reeves, Thompson and Richardson	Proactive composite alliance	131800	147936	112	successful	1396	105.97	US	USD	1406523600	1406523600	7/28/14	7/28/14	FALSE	FALSE	film & video/drama	film & video	drama

Using the Excel workbook in your .zip file, modify and analyze the sample-project data and try to uncover market trends.

- Data for this dataset was generated by edX Boot Camps LLC, and is intended for educational purposes only.
- Use conditional formatting to fill each cell in the `outcome` column with a different color, depending on whether the associated campaign was successful, failed, canceled, or is currently live.
  - Create a new column called `Percent Funded` that uses a formula to find how much money a campaign made relative to its initial funding goal.

- Use conditional formatting to fill each cell in the `Percent Funded` column according to a three-color scale. The scale should start at 0 with a dark shade of red, and it should transition to green at 100 and blue at 200.
  - Create a new column called `Average Donation` that uses a formula to find how much each project backer paid on average.
  - Create two new columns, one called `Parent Category` and another called `Sub-Category`, that use formulas to split the `Category and Sub-Category` column into the two new, separate columns.

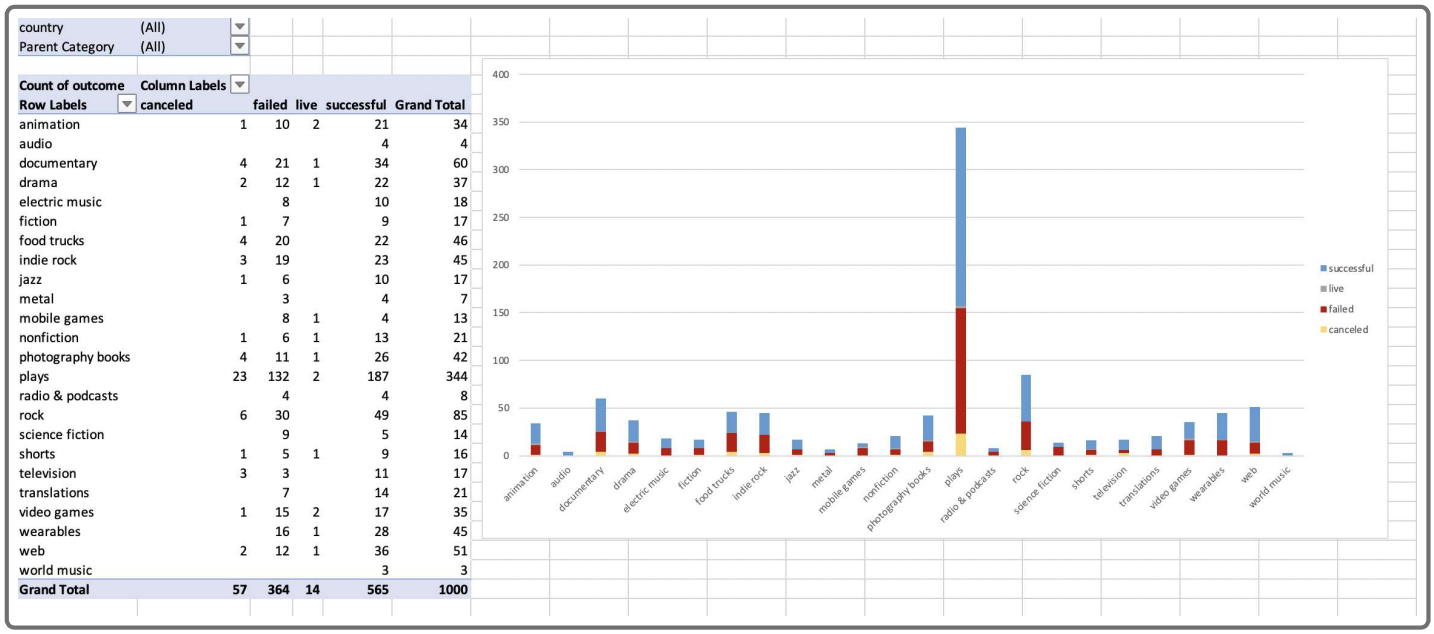


country (All)

Count of outcome	Column Labels				
Row Labels	canceled	failed	live	successful	Grand Total
film & video	11	60	5	102	178
food	4	20		22	46
games	1	23	3	21	48
journalism				4	4
music	10	66		99	175
photography	4	11	1	26	42
publishing	2	24	1	40	67
technology	2	28	2	64	96
theater	23	132	2	187	344
Grand Total	57	364	14	565	1000

- Create a new sheet with a pivot table that analyzes your initial worksheet to count how many campaigns were successful, failed, canceled, or are currently live per **category**.

- Create a stacked-column pivot chart that can be filtered by country based on the table that you created.



country	(All)					
Parent Category	(All)					
Count of outcome	Column Labels					
Row Labels	canceled	failed	live	successful	Grand Total	
animation	1	10	2	21	34	
audio				4	4	
documentary	4	21	1	34	60	
drama	2	12	1	22	37	
electric music		8		10	18	
fiction	1	7		9	17	
food trucks	4	20		22	46	
indie rock	3	19		23	45	
jazz	1	6		10	17	
metal		3		4	7	
mobile games		8	1	4	13	
nonfiction	1	6	1	13	21	
photography books	4	11	1	26	42	
plays	23	132	2	187	344	
radio & podcasts		4		4	8	
rock	6	30		49	85	
science fiction		9		5	14	
shorts	1	5	1	9	16	
television	3	3		11	17	
translations		7		14	21	
video games	1	15	2	17	35	
wearables		16	1	28	45	
web	2	12	1	36	51	
world music				3	3	
Grand Total	57	364	14	565	1000	

- Create a new sheet with a pivot table that analyzes your initial sheet to count how many campaigns were successful, failed, or canceled, or are currently live per **sub-category**.

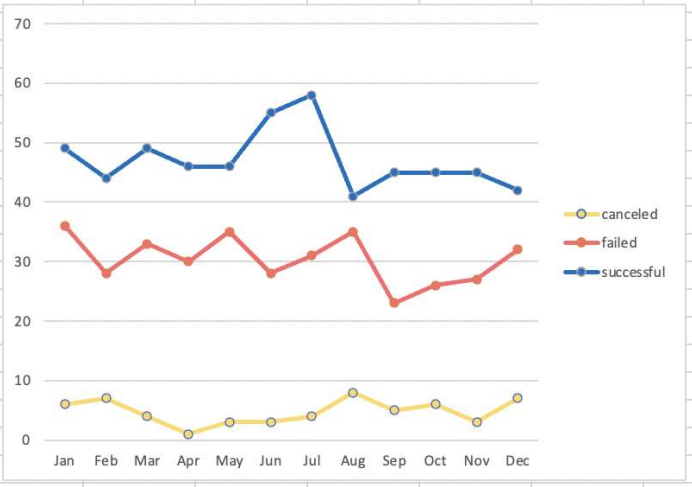
- Create a stacked-column pivot chart that can be filtered by country and parent category based on the table that you created.

- The dates in the `deadline` and `launched_at` columns use Unix timestamps. Fortunately for us, **this formula (https://www.extendoffice.com/documents/excel/2473-excel-timestamp-to-date.html)** that can be used to convert these timestamps to a normal date.

    - Create a new column named `Date Created Conversion` that will use **this formula (https://www.extendoffice.com/documents/excel/2473-excel-timestamp-to-date.html)** to convert the data contained in `launched_at` into Excel's date format.

    - Create a new column named `Date Ended Conversion` that will use **this formula (https://www.extendoffice.com/documents/excel/2473-excel-timestamp-to-date.html)** to convert the data contained in `deadline` into Excel's date format.

Parent Category	(All)			
Years	(All)			
Count of outcome	Column Labels			
Row Labels	canceled	failed	successful	Grand Total
Jan	6	36	49	91
Feb	7	28	44	79
Mar	4	33	49	86
Apr	1	30	46	77
May	3	35	46	84
Jun	3	28	55	86
Jul	4	31	58	93
Aug	8	35	41	84
Sep	5	23	45	73
Oct	6	26	45	77
Nov	3	27	45	75
Dec	7	32	42	81
Grand Total	57	364	565	986



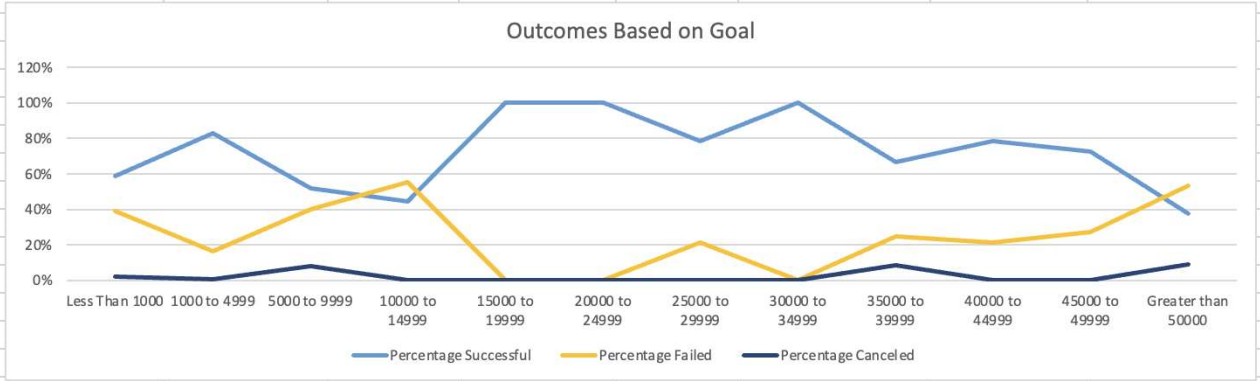
  - Create a new sheet with a pivot table that has a column of `outcome`, rows of `Date Created Conversion`, values based on the count of `outcome`, and filters based on `parent category` and `Years`.
  - Now, create a pivot-chart line graph that visualizes this new table.
- Create a report in Microsoft Word, and answer the following questions:
  - Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  - What are some limitations of this dataset?
  - What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

Crowdfunding Goal Analysis

- Create a new sheet with 8 columns:
  - `Goal`
  - `Number Successful`
  - `Number Failed`
  - `Number Canceled`
  - `Total Projects`
  - `Percentage Successful`
  - `Percentage Failed`
  - `Percentage Canceled`
- In the `Goal` column, create 12 rows with the following headers:
  - Less than 1000
  - 1000 to 4999
  - 5000 to 9999

- 10000 to 14999
- 15000 to 19999
- 20000 to 24999
- 25000 to 29999
- 30000 to 34999
- 35000 to 39999
- 40000 to 44999
- 45000 to 49999
- Greater than or equal to 50000

Goal	Number Successful	Number Failed	Number Canceled	Total Projects	Percentage Successful	Percentage Failed	Percentage Canceled
Less Than 1000	30	20	1	51	59%	39%	2%
1000 to 4999	191	38	2	231	83%	16%	1%
5000 to 9999	164	126	25	315	52%	40%	8%
10000 to 14999	4	5	0	9	44%	56%	0%
15000 to 19999	10	0	0	10	100%	0%	0%
20000 to 24999	7	0	0	7	100%	0%	0%
25000 to 29999	11	3	0	14	79%	21%	0%
30000 to 34999	7	0	0	7	100%	0%	0%
35000 to 39999	8	3	1	12	67%	25%	8%
40000 to 44999	11	3	0	14	79%	21%	0%
45000 to 49999	8	3	0	11	73%	27%	0%
Greater than 50000	114	163	28	305	37%	53%	9%



- Using the `COUNTIFS()` formula, count how many successful, failed, and canceled projects were created with goals within the ranges listed above. Populate the `Number Successful`, `Number Failed`, and `Number Canceled` columns with these data points.
- Add up each of the values in the `Number Successful`, `Number Failed`, and `Number Canceled` columns to populate the `Total Projects` column. Then, using a mathematical formula, find the percentage of projects that were successful, failed, or canceled per goal range.
- Create a line chart that graphs the relationship between a goal amount and its chances of success, failure, or cancellation.

Statistical Analysis

Most people would use the number of campaign backers to assess the success of a crowdfunding campaign. Creating a summary statistics table is one of the most efficient ways that data scientists can characterize quantitative metrics, such as the number of campaign backers.

For gaining an in-depth understanding of campaign backers, evaluate the number of backers of successful and unsuccessful campaigns by creating **your own** summary statistics table.

- Create a new worksheet in your workbook, and create one column for the number of backers of successful campaigns and one column for unsuccessful campaigns.

outcome	backers_count		outcome	backers_count
successful	158		failed	0
successful	1425		failed	24
successful	174		failed	53
successful	227		failed	18
successful	220		failed	44
successful	98		failed	27
successful	100		failed	55
successful	1249		failed	200
successful	1396		failed	452

- Use Excel to evaluate the following values for successful campaigns, and then do the same for unsuccessful campaigns:
    - The mean number of backers
    - The median number of backers
    - The minimum number of backers
    - The maximum number of backers
    - The variance of the number of backers
    - The standard deviation of the number of backers

- Use your data to determine whether the mean or the median better summarizes the data.
- Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

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 apivot 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 to your Dropbox or Google Drive folder 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.