



MSBX 5420 Group Project Presentation



Group 9:
Bella Franz, Danielle Allen, Morgan Likens, Rachel
Raifsnider, Samantha Fildish




Dataset Overview -

- Here is a quick view of how our dataset was presented as a csv file
- It was sourced from Kickstarter (<https://www.kaggle.com/datasets/kemical/kickstarter-projects?select=ks-projects-201801.csv>)
- There are 378,661 rows of data for projects ranging from 1970 - 2018
- We did have to do some data pre processing as well as a custom schema that we used to read in the file

ID	name	category	main_category	currency	deadline	goal	launched	pledged	state	backers	country	usd pledged	usd_pledged_real	usd_goal_real
1000002330	The Songs of Adelaide & Abullah	Poetry	Publishing	GBP	10/9/15	1000	8/11/15	0	failed	0	GB	0	0	1533.95
1000003930	Greeting From Earth: ZGAC Arts Capsule For ET	Narrative Film	Film & Video	USD	11/1/17	30000	9/2/17	2421	failed	15	US	100	2421	30000
1000004038	Where is Hank?	Narrative Film	Film & Video	USD	2/26/13	45000	1/12/13	220	failed	3	US	220	220	45000
1000007540	ToshiCapital Rekordz Needs Help to Complete Album	Music	Music	USD	4/16/12	5000	3/17/12	1	failed	1	US	1	1	5000
1000011046	Community Film Project: The Art of Neighborhood Filmmaking	Film & Video	Film & Video	USD	8/29/15	19500	7/4/15	1283	canceled	14	US	1283	1283	19500
1000014025	Monarch Espresso Bar	Restaurants	Food	USD	4/1/16	50000	2/26/16	52375	successful	224	US	52375	52375	50000
1000023410	Support Solar Roasted Coffee & Green Energy! SolarCoffee.co	Food	Food	USD	12/21/14	1000	12/1/14	1205	successful	16	US	1205	1205	1000
1000030581	Chaser Strips. Our Strips make Shots their B*tch!	Drinks	Food	USD	3/17/16	25000	2/1/16	453	failed	40	US	453	453	25000
1000034518	SPIN - Premium Retractable In-Ear Headphones with Mic	Product Design	Design	USD	5/29/14	125000	4/24/14	8233	canceled	58	US	8233	8233	125000
1000041195	STUDIO IN THE SKY - A Documentary Feature Film (Canceled)	Documentary	Film & Video	USD	8/10/14	65000	7/11/14	6240.57	canceled	43	US	6240.57	6240.57	65000
100004721	Of Jesus and Madmen	Nonfiction	Publishing	CAD	10/9/13	2500	9/9/13	0	failed	0	CA	0	0	2406.39
100005484	Lisa Lim New CD!	Indie Rock	Music	USD	4/8/13	12500	3/9/13	12700	successful	100	US	12700	12700	12500
1000055792	The Cottage Market	Crafts	Crafts	USD	10/2/14	5000	9/2/14	0	failed	0	US	0	0	5000
1000056157	G-Spot Place for Gamers to connect with eachother & go prol	Games	Games	USD	3/25/16	200000	2/9/16	0	failed	0	US	0	0	200000
1000057089	Tombstone: Old West tabletop game and miniatures in 32mm.	Tabletop Games	Games	GBP	5/3/17	5000	4/5/17	94175	successful	761	GB	57763.78	121857.33	6469.73
1000064368	Survival Rings	Design	Design	USD	2/28/15	2500	1/29/15	664	failed	11	US	664	664	2500
1000064918	The Beard	Comic Books	Comics	USD	11/8/14	1500	10/9/14	395	failed	16	US	395	395	1500
1000068480	Notes From London: Above & Below	Art Books	Publishing	USD	5/10/15	3000	4/10/15	789	failed	20	US	789	789	3000
1000070642	Mike Corey's Darkness & Light Album	Music	Music	USD	8/17/12	250	8/2/12	250	successful	7	US	250	250	250
1000071625	Boco Tea	Food	Food	USD	6/2/12	5000	5/3/12	1781	failed	40	US	1781	1781	5000
1000072011	CMUK. Shoes: Take on Life Feet First.	Fashion	Fashion	USD	12/30/13	20000	11/25/13	34268	successful	624	US	34268	34268	20000
1000081649	Mikey! clothing brand fundraiser	Childrenswear	Fashion	AUD	9/7/17	2500	8/8/17	1	failed	1	AU	0	0.81	2026.1
1000082254	Alice in Wonderland in G Minor	Theater	Theater	USD	6/15/14	3500	5/16/14	650	failed	12	US	650	650	3500

Kickstarter -

- Kickstarter was started to help creators and creatives connect directly with the community
- The goal is to help bring creator projects to life with funding opportunities for every day people that give unique rewards based on how much you pledge to the project
- Some of the most successful projects (monetarily) of all time on Kickstarter include:
 - “Four Secret Novels” by Brandon Sanderson - \$41 million pledged
 - “Pebble Time” Smartwatch - \$20 million pledged
 - “Frosthaven” table game - \$13 million pledged



Pebble Time - Awesome Smartwatch, No Compromises

\$20,338,986
pledged of \$500,000 goal

78,471
backers

Project We Love Palo Alto, CA Product Design

Support

Pledge US\$ 179 or more

Your choice of one Pebble Time watch in any of the three colors. Regular retail price will be \$199.

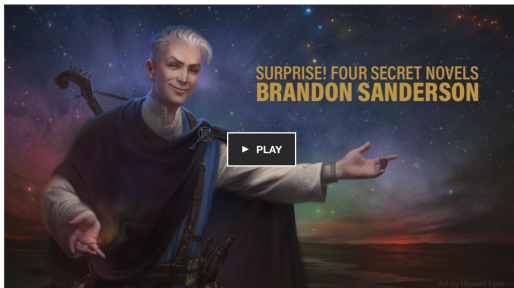
ESTIMATED DELIVERY May 2015 SHIPS TO Anywhere in the world

22,991 backers
(Limited @ left of 23000)

Pledge \$179 or more

Your choice of one Pebble Time watch in any of the three colors. Regular retail price will be \$199.

ESTIMATED DELIVERY Jun 2015 SHIPS TO Anywhere in the world



Surprise! Four Secret Novels by Brandon Sanderson

\$41,754,153
pledged of \$1,000,000 goal

185,341
backers

Project We Love Pleasant Grove, UT Fiction

Support

Pledge \$40 or more

Quarterly Books - Ebook

A new ebook available for download in January, April, July, and October of 2023.

INCLUDES:

- Quarterly Ebooks

ESTIMATED DELIVERY Jan 2023

44,430 backers Add-ons

Pledge US\$ 60 or more

Quarterly Books - Audiobook

A new audiobook available for download in January, April, July, and October of 2023.

Project overview/ Business Questions we are Trying to Answer

- Which kickstarter categories and sub categories have the highest rate of success?
- How does the rate of success by category change over time?
- Are there any other factors that impact the success of a project?
- Utilized general business analytics to help get the details we need to answer our main questions for the dataset

Analysis/ statistics/ methods

- Methods used for analysis
 - To read our data accurately, we had to create a custom Schema and apply that schema during reading
 - To analyze our data, we used a combination of dataframe operations and Spark SQL
 - We also converted some dataframes to Pandas Dataframes to create better visualizations

Analysis/ Most Successful Main Categories

We have identified the Main Category with the highest success rate as Dance, with a 62% success rate.

While Technology has the lowest success rate at 19%

main_category	total_count	total_success	success_rate
Dance	3768	2338	0.6204883227176221
Theater	10913	6534	0.5987354531292953
Comics	10819	5842	0.5399759682040854
Music	51918	24197	0.466061866789938
Art	28153	11510	0.408837424075587
Film & Video	63585	23623	0.3715184398836204
Games	35231	12518	0.35531208310862594
Design	30070	10550	0.35084802128367143
Publishing	39874	12300	0.30847168581030243
Photography	10779	3305	0.3066147137953428
Food	24602	6085	0.24733761482806277
Fashion	22816	5593	0.2451349929873773
Crafts	8809	2115	0.24009535702122828
Journalism	4755	1012	0.2128286014721346
Technology	32569	6434	0.19754981731093985

Analysis/ Most Successful Sub-categories

When broken down further into sub categories, the sub-category with the highest success rate is Chiptune Music, with a 77% success rate. However, we can see that the total number of projects is not very high.

main_category	category	total_count	total_success	success_rate
Music	Chiptune	35	27	0.7714285714285715
Comics	Anthologies	405	303	0.7481481481481481
Dance	Residencies	69	50	0.7246376811594203
Dance	Dance	2322	1542	0.6640826873385013
Crafts	Letterpress	38	25	0.6578947368421053

Analysis/ Least Successful Sub-categories

The sub-category with the lowest success rate is App - Technology, with a 6% success rate, followed closely by Web and Mobile Games.

main_category	category	total_count	total_success	success_rate
Technology	Apps	6345	378	0.059574468085106386
Technology	Web	3906	258	0.06605222734254992
Games	Mobile Games	1789	153	0.08552263834544438
Journalism	Video	428	51	0.1191588785046729
Technology	Software	3048	371	0.12171916010498687

Analysis/ A Categories Success Rate Over the Years

- For this example we pulled a random main category (Design) to review the success rate from 1970 to 2018
- Overall its average success rate was 28.89%, the best year was 2010 with a success rate of 41.57%
- While the success rate generally seems to stay in a window from 30% to 41% the total count & the success counts show a wide variety by year showing its pretty consistent

	main_category	year	total_count	successful_count	success_rate
0	Design	1970	1	0	0.000000
1	Design	2009	31	8	0.258065
2	Design	2010	166	69	0.415663
3	Design	2011	903	302	0.334441
4	Design	2012	1841	760	0.412819
5	Design	2013	3246	1255	0.386630
6	Design	2014	5394	1718	0.318502
7	Design	2015	6705	2046	0.305145
8	Design	2016	5908	2131	0.360697
9	Design	2017	5866	2261	0.385442
10	Design	2018	9	0	0.000000

Analysis/ Average Funding for a Project

As you can imagine, projects in certain categories require more funding than others. We found the average Goal and Amount Pledged per Main category.

- Projects labeled under Technology have the highest average goal at \$102,289
 - While only having an average pledge amount of \$21,151
- Projects labeled under Design have the highest average pledge amount of \$24,417
 - While only having an average goal amount of \$41,871
 - This seems more closely inline with the amount of pledge they receive.

main_category	avg(usd_pledged_real)	avg(usd_goal_real)
Technology	21151.171184047813	102288.54276458472
Journalism	2615.702569326398	86003.39486431722
Film & Video	6157.978426427467	82625.2922127824
Food	5113.669713271148	49039.62093328174
Games	21041.90315417259	44923.74139177793
Design	24417.034673355178	41871.38861747499
Art	3220.986968480684	39198.576013676124
Theater	4005.918097030627	27088.321180519008
Publishing	3349.951929046906	24798.240252090338
Fashion	5712.450837448087	22412.409349196747
Comics	6610.443830018322	19571.430671694034
Music	3858.367606525988	15372.161812904633
Photography	3571.9093264821718	12266.843695389836
Crafts	1632.907281949969	10434.466230333075
Dance	3452.94635507167	9588.328559136442

Analysis/ Does funding mean Success?

One question we asked was, if a Project is fully funded, meaning the pledge amount met its goal, is it more likely to be successful?

- The answer seemed to lean towards Yes.
 - Out of all of our projects that met or exceeded their goal,
 - The majority of them are “successful”.

state	projects_in_state
successful	133951
undefined	1654
canceled	698
live	441
suspended	292
failed	6

Analysis/Funding Amounts for Main Categories

Looking into the largest pledged amount and total pledged for all kickstarter projects based on their main category, offers some interesting comparisons

- To start, over 2 billion dollars in total was pledged
- Games received the most funds int total with 574,402,021
- The highest funded project went to Design for 20,338,986
- Design was also the second highest in overall pledge amount
- The Dance category's largest funded project only 146,075

main_category	max pledged	total pledge
Games	8782572.0	5.744020215326892E8
Design	2.0338986E7	5.1213200917237395E8
Technology	6225355.0	5.0332082952915287E8
Film & Video	5764229.5	3.468760309230262E8
Music	1192793.1	1.662657910970528E8
Publishing	675614.06	1.0371388403976554E8
Food	1811321.6	1.0356017260043496E8
Fashion	1105177.5	9.762922885520282E7
Art	1226811.5	7.13413098417696E7
Comics	1254120.1	5.781691844161123E7
Theater	318120.34	3.8785122693737805E7
Photography	683801.4	3.017172002512759E7
Dance	146075.97	1.1161583888581932E7
Crafts	204127.52	1.0387490141041458E7
Journalism	194635.66	1.0252217015961766E7

Analysis/Funding Amounts Cont.

- Here, we can see what percent of the total amount pledged for all projects went to each main

main_category	total pledged	percent
Journalism	1.0252217015961766E7	0.003886630354111373
Crafts	1.0387490141041458E7	0.003937912592207941
Dance	1.1161583888581932E7	0.004231372655668...
Photography	3.017172002512759E7	0.011438142862449934
Theater	3.8785122693737805E7	0.014703496318378887
Comics	5.781691844161123E7	0.021918477715258095
Art	7.13413098417696E7	0.02704559412870316
Fashion	9.762922885520282E7	0.03701138238942369
Food	1.0356017260043496E8	0.03925981177331766
Publishing	1.0371388403976554E8	0.03931808400311403
Music	1.662657910970528E8	0.06303160277645803
Film & Video	3.468760309230262E8	0.13150120689018943
Technology	5.0332082952915287E8	0.19080965715599488
Design	5.1213200917237395E8	0.1941499881501158
Games	5.744020215326892E8	0.21775664023460523

- Overall, the majority of funds were distributed to mainly 4 categories which comprised of approx. 73% of all pledged to kickstarter
- The most significant portions went to Design and Games
- Design accounted for 19.4% and Games achieved 21.78%
- On the lower end, we have journalism gaining a mere 0.39%

Analysis/ Launch and Deadline Month Impact on Success

Outside of December and January the launch month and deadline month have a very small impact on the success of a project

Launch Month-

- March has the highest rate of successful projects but there is not a significant change in the top 11 months all ranging from 38% to 32%
 - ~6% difference between March and second lowest month (July)
- December has the lowest rate of success with a drop of about 8%

Launch Month

Month	total_successful_projects	total_count	success_rate
3	12851	33946	0.37857184940788313
4	12093	32260	0.3748605083694978
2	11003	29732	0.3700726489977129
10	12350	33490	0.3687667960585249
5	12062	32888	0.36675991243006567
9	11231	31069	0.3614857253210596
6	11773	32623	0.3608803604818686
11	11770	32890	0.35785953177257523
1	9480	27841	0.3405050105958838
8	10820	32291	0.33507788547892603
7	11681	36367	0.3211977892044986
12	6842	23264	0.29410247592847316

Deadline

- April has the highest rate of successful projects but again we do not see a significant change in the top 11 months all ranging from 38% to 33%
 - ~5% difference between April and second lowest month (August)
- January has the lowest rate of success with a drop of about 9%

Deadline Month

Month	total_successful_projects	total_count	success_rate
4	12344	32566	0.3790456304120862
5	12604	33759	0.3733522912408543
3	12137	32574	0.3725977773684534
11	11801	31990	0.36889653016567675
6	11711	32050	0.3653978159126365
7	12103	33703	0.3591074978488562
10	11467	32032	0.3579857642357642
12	11799	33190	0.3554986441699307
2	8358	24485	0.3413518480702471
9	10551	31760	0.3322103274559194
8	11959	36289	0.32954889911543445
1	7122	24263	0.2935333635576804

Analysis/ Geographic Impacts on Success

Geographic location of the projects has a small impact on the success rate

- The US has the highest total number of successful projects
 - The US has about 9 times as many total projects on the platform than the next highest country (Great Britain)
- The US also has the highest success rate
 - Even with the much greater volume of project the success rate is still only 1.5% higher than the next highest country (Great Britain)
- Among the top 10 countries the success rate remains between 28% and 37%

country	total_successful_projects	total_count	success_rate
US	109299	292627	0.3735096214635013
GB	12067	33672	0.3583689712520789
HK	216	618	0.34951456310679613
DK	360	1113	0.32345013477088946
SG	178	555	0.3207207207207207
NZ	448	1447	0.30960608154803043
FR	908	2939	0.3089486219802654
LU	19	62	0.3064516129032258
SE	509	1757	0.2896983494593056
CA	4134	14756	0.28015722417999456

Analysis/ Rate of Successful Crowd Funding Over Time

- What percent of projects met their crowdfunding goal?
 - The entertainment sector (Dance and Theater) had the highest rate of crowdfunding success over time
 - The top 25 categories had a success rate of over 50% of their funding goal

Year	main_category	total_successful_crowdfunding	total_count	success_rate
2011	Dance	289	400	0.7225
2013	Dance	385	536	0.7182835820895522
2012	Dance	379	528	0.7178030303030303
2012	Theater	1186	1809	0.6556108347153123
2013	Theater	973	1489	0.6534586971121558
2011	Theater	926	1425	0.6498245614035087
2010	Dance	89	137	0.6496350364963503
2017	Comics	1270	1997	0.6359539308963446
2014	Dance	416	699	0.5951359084406295
2016	Dance	261	439	0.5945330296127562
2010	Theater	325	551	0.5898366606170599
2016	Theater	674	1150	0.5860869565217391
2016	Comics	1091	1874	0.5821771611526148
2017	Dance	219	378	0.5793650793650794
2011	Music	3636	6348	0.5727788279773157
2014	Theater	1021	1793	0.5694366982710541
2009	Comics	9	16	0.5625
2017	Theater	506	905	0.5591160220994476
2012	Music	5050	9066	0.557026251930289
2013	Music	3843	6984	0.5502577319587629
2013	Comics	746	1396	0.5343839541547278
2015	Theater	904	1701	0.5314520870076426
2014	Comics	824	1569	0.5251752708731676
2010	Music	1096	2087	0.5251557259223766
2011	Art	1157	2229	0.5190668461193361
2015	Comics	996	1921	0.5184799583550235
2009	Music	97	189	0.5132275132275133
2013	Art	1676	3409	0.4916397770607216

Analysis/ Which projects over shot their crowdfunding goal?

- Calculated this by the amount funded divided by their goal
- The music category took the top 3 spots when analyzing which projects overshoot their crowdfunding goal
- VULFPECK overshoot their crowdfunding goal by 100,000 times!

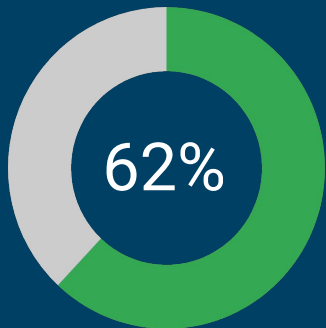
main_category	name	max(overshoot)
Music	VULFPECK /// The ...	104277.890625
Music	Re-covering with ...	68764.1015625
Music	VULFPECK /// Thri...	55266.5703125
Games	Energy Hook	41535.01171875
Games	Band of Brothers ...	32843.0
Art	CLOCKWRIGHT: Larg...	27588.23046875
Music	Heartbreak Hits -...	24675.0
Publishing	"Penny Arcade's P...	23036.0671875
Technology	THE 'mi8' RISES ...	22603.0
Games	CO-OP THE GAME (C...	16970.855652793736
Games	Worst Game Ever	15804.0
Film & Video	Y (A Generation)	15066.0
Music	Graveface Archiva...	14328.76953125
Games	Braiiiins! ...it ...	12984.0
Crafts	Multi-Purpose Al...	12575.150390625
Music	Joe Hertler & The...	12497.26953125
Technology	1Ena - Movement t...	11644.1103515625
Games	d6xd6 CORE Role-P...	11456.3603515625
Art	LOVELAND Round 6:...	10000.000223517423
Technology	Arduino IoT. Home...	9605.0

only showing top 20 rows

Initial findings/ Success of Categories

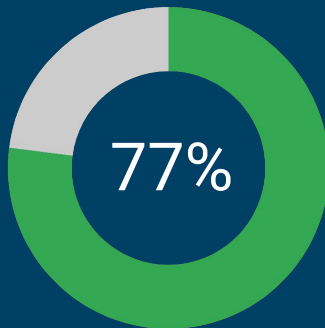
Category with the highest success rate:

Dance



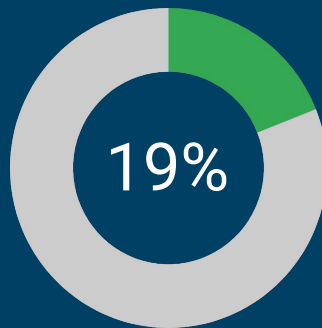
Sub-category with the highest success rate:

Chiptune Music



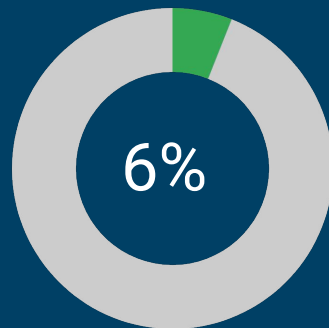
Category with lowest success rate:

Technology



Sub-Category with the lowest success rate:

App- Technology



Initial Findings/ Success Rate Over the Years

- While some individual years saw larger changes the trends for all categories remained within +/- 4%
 - Negative value = average downward trend, positive value = average upward trend, 0 = neutral trend
- Dance, Theater, Design, Fashion and Comics saw small positive trends (1-2%)
- The other categories all saw small negative trends (1-4%)
- Note: We dropped the years 1970 and 2018, as neither of these years had values and were impacting the truthfulness of the trend

year	2009	2010	2011	2012	2013	2014	2015	2016	2017	trend
main_category										
Dance	0.375000	0.639053	0.723971	0.721569	0.717190	0.586648	0.470305	0.609302	0.551351	0.022044
Theater	0.410714	0.591433	0.654657	0.654300	0.652881	0.560732	0.535036	0.586572	0.538206	0.015936
Design	0.258065	0.415663	0.334441	0.412819	0.386630	0.318502	0.305145	0.360697	0.385442	0.015922
Fashion	0.153846	0.248555	0.308458	0.264117	0.321992	0.225294	0.207315	0.224594	0.273397	0.014944
Comics	0.521739	0.413534	0.454237	0.470386	0.539398	0.518331	0.517170	0.586631	0.619381	0.012205
Art	0.455621	0.446262	0.508120	0.493446	0.489535	0.353489	0.320804	0.366099	0.426076	-0.003693
Film & Video	0.376900	0.399729	0.414470	0.401256	0.422549	0.361491	0.294452	0.345512	0.345365	-0.003942
Games	0.500000	0.299145	0.354557	0.338246	0.364303	0.330226	0.305069	0.373379	0.416510	-0.010436
Photography	0.424658	0.407725	0.380460	0.343750	0.357546	0.228535	0.246119	0.350000	0.338787	-0.010734
Music	0.532203	0.525142	0.568859	0.558014	0.546377	0.414062	0.305380	0.416782	0.434630	-0.012197
Publishing	0.441860	0.314894	0.335993	0.306412	0.336263	0.284130	0.263700	0.326390	0.342786	-0.012384
Crafts	0.363636	0.345455	0.393162	0.398734	0.440994	0.198231	0.203324	0.226170	0.258755	-0.013110
Technology	0.316667	0.288889	0.306346	0.399754	0.368018	0.177045	0.164967	0.185917	0.194120	-0.015318
Journalism	0.487805	0.361702	0.394118	0.316547	0.387560	0.184946	0.168478	0.172619	0.197432	-0.036297
Food	0.531250	0.465116	0.410828	0.376796	0.419212	0.214330	0.190036	0.208487	0.223684	-0.038446

Business implications

- Using our analysis we can determine future projects that may be more likely to succeed as well as projects that are not likely to do well on Kickstarter
- An individual, as well as a business, could also use this to determine if Kickstarter is the right platform for their project
 - If it is the right platform then you could customize your project to be more likely to succeed given the analytics available