
A STATISTICAL ANALYSIS OF TWITTER POSTS TO IMPROVE CANCER COUNCIL NSW'S SOCIAL MEDIA STRATEGY

Gavan Corke

Contents

INTRODUCTION	2
RATIONALE AND SITUATION ANALYSIS	2
TWITTER.....	3
FRAMEWORKS	4
R.O.P.E.S	4
RESEARCH QUESTIONS	6
METHODOLOGY and ANALYSIS.....	7
ASSUMPTIONS	7
METHODOLOGY	7
SAMPLES, DATA COLLECTION and ANALYSIS.....	8
RESULTS	10
QUESTION 1	10
QUESTION 2	12
QUESTION 3	14
QUESTION 4	14
DISCUSSION and RECOMMENDATIONS.....	16
REFERENCES.....	19
APPENDIXES.....	22

INTRODUCTION

The aim of this report is to improve the overall awareness and donations for the Cancer Council NSW via its social media strategy on twitter. This will be done by performing a content analysis of Twitter posts (or tweets) for 10 Australian charities to answer a set of research questions which ascertain to the goal of improving the content of twitter posts (measured by the number of favorites or likes) from CCNSW which will hope to increase awareness and donations in the short and long term.

RATIONALE AND SITUATION ANALYSIS

Charities play an important role in society, especially within Australia. They provide a wide range of services for the community that may otherwise not be fulfilled in the for-profit system. On a social level, they provide services for communities in need, whilst on an economic level they provide employment for over 1 million Australians (Johns, 2020). The revenue of charities in Australia in 2018 accounted for over 8% of national GDP (Social Ventures Australia and the Centre for Social Impact, 2020, p 6).¹ But most importantly, on a moral and ethical level, charities represent and reflect the collective effort of people to help others less fortunate, or in extenuating conditions. Therefore, the collective importance of such organizations in providing services for communities and individuals for the goal of improving society cannot be underestimated.

During the current COVID-19 pandemic, many Australian charities have been hit negatively. The SVA report states that cancellation of services has occurred for various reasons such as social distancing measures which ultimately affect the ability of charities to generate income (in the forms of donations and fundraising etc.). Ranging from a lack of live events and social gatherings, to income shocks that have affected the ability to pay staff, and loss of volunteers are some of the issues caused or exacerbated by the current pandemic (SVA, p 12). One statistic states that 78% of the 170 charities surveyed reported a down-turn in revenue due to COVID 19. (SVA, p 12). This represents a systemic issue facing the charity industry within Australia. The loss of income has come from various sources contributing less than previously. On the fundraising side, 67% of charities from March and April last year reported a decrease in donations (SVA, p12). The government response has been to provide extra funding to help mitigate some of the losses as can be seen in the SVA report (SVA, p 14).

The problem with this is the double bind effect the pandemic has had on charities. On the one hand, charities have been unable to successfully generate enough revenue as they had been able to previously due to the pandemic, but also due to the pandemic, services and the resources that they provide have in general been in greater demand, putting charities in this double bind problem (SVA, p 5). The economic consequences of Covid 19 on charities and the broader Australia economy have been immediate, causing charities to have a large reduction in income. Governmental response to the crisis by implementing social distancing and lockdown laws on a state-by-state basis has affected the ability of charities to effectively navigate the change in economic landscape with any certainty and consistency.

The covid and charities report by SVA has projected a 20% overall drop in income for Australian charities collectively. The report estimates that after six months 1700 charities would have run out of current assets (net) to cover their deficits, which is approximately 17% of the charities surveyed. (SVA, p 15) 71% would be operating at a deficit but with enough current assets (net) to cover their deficits over the next 6 months. And lastly, only the remaining 12% of charities would be able to survive with an operating surplus over the projected 6-month period starting from April 2020. (SVA, p 15).

This information on charities in Australia and the economic and financial impact covid-19 has had on their overall health provides the context and backdrop from which this report will develop its analysis and recommendations. In this next section, we will look at the Cancer Council NSW (CCNSW henceforth) specifically and how it has been affected by Covid-19 and how this report aims to help the CCNSW.

The Cancer Council Australia commenced in 1961 to raise awareness about cancer related issues on a national level. The Cancer Council NSW is a non-profit, charitable organization, which deals with cancer awareness,

¹ Henceforth SVA

research and donations at the state level for NSW (Cancer Council Australia | Our History, 2021). Cancer Council NSW currently has 254 full time staff, 101 part time staff with an annual revenue of \$73 million in 2018 (Change Path | Cancer Council NSW - organization details, 2021). It is a member of the Cancer Council Federation which comprises of all the different state subsidiaries of The Cancer Council with CCNSW performing its operations in NSW. It is important to note that the Cancer Council Federation has been affected by COVID-19 negatively. In the Cancer Council Australia's annual report for 2019/2020 they state that the COVID-19 crisis has an 'unprecedented effect' on the Australian health system and the current strategy that the Cancer Council Australia had planned to execute for until the end of 2021 must be now 'put on hold' (Cancer Council Australia | Annual Review, p 31, 2020)

Looking at CCNSW specifically, the Australian bushfires in early 2020 and the COVID-19 pandemic effected the financial status of the charity adversely.

In 2019/20, our **fundraising income was significantly impacted**, first by the bushfires, and then by the COVID-19 pandemic. We are facing an almost 30% decline in fundraising income over the next year, as we have had to stop, postpone or adapt much of our fundraising activity.

We are so grateful to our wonderful supporters who continued to fundraise and donate to us, including those who supported our emergency appeal, which raised more than **\$1.2 million**, making it our most successful appeal to date.



	2016	2017	2018	2019	2020
Total fundraising income (\$m)	67.4	66.9	56.5	63.8	48.1
Number of people who left us a gift in their will	152	150	108	147	134

FIGURE 1: CANCER COUNCIL NSW TOTAL FUNDRAISING REVENUE INFOGRAPHIC. (CANCER COUNCIL NSW | FINANCIAL REPORT 2019/2020, 2020)

As can be seen in figure 1, the total fundraising in 2020 reached a total of 48.1 million AUD. This is in contrast to the year before of 63.8 million AUD. This equates to a 15.7-million-dollar difference or a 25% reduction from the previous year (2019). This is quite a substantial amount of fundraising money that is lost due to the pandemic. The CCNSW states that another 30% decline in fundraising income should be expected for 2021. Therefore, it is paramount for the organization that it can extract the most out of its current fundraising strategies and campaigns to increase its total fundraising income to above 2020 levels or at the very least, reduce the level of its decline for the subsequent year.

The aim of this report is to improve CCNSW social media strategy on twitter so that it can offset the 30% decline in fundraising that is projected for 2021. In the next section we will justify improving CCNSW twitter social media campaign strategy to improve awareness and subsequently donations and what that entails methodologically.

TWITTER

The SVA report states that one area that many charities have been underutilizing is the 'online' space for their operations, as only 34% of charities have an online presence and 55% state that a lack of funding has hampered their ability to utilize technology effectively (SVA, p5). The SVA recommends that charities utilize online platforms more to help with its activities in the post-pandemic world (SVA, p 25). While this is not necessarily a problem for the CCNSW, as they do have a large online presence, this however speaks to the importance that having an online presence can have for a charitable organization.

One popular avenue for charities to utilize in the online space is social media. Microblogging sites are mediums which charitable organizations have been using to perform some of their objectives such as for raising awareness, reporting on events related to the organization or charity, and receiving donations (Social Misfits

Media, 2014, p 2). Twitter is one example of a microblogging site. In twitter, users post brief messages which are called 'tweets' which comprise of a maximum of 280 characters (Svensson et al., 2014, p 4). These can be combined with hashtags and images to make the post easier to identify and to improve the overall aesthetic of the message. It is also possible to retweet a particular post and to 'favorite' or 'like' posts. Unfortunately, many charities do not utilize these microblogging sites (facebook, twitter, Instagram etc.) effectively. For example, charities do not always regularly employ hashtags which can help them circumvent the character limit on a tweet (Svensson et al., p 4). A survey found that charities that embraced the open and transparent nature of social media reported increased volunteering and donations (Svensson et al., page 5). Therefore, utilizing social media can be one part of an effective overall strategy to increase awareness and possibly donations for a charity or non-profitable organization. By improving its Twitter utilization, the CCNSW should be able to improve its awareness and donations on some objective measure.

FRAMEWORKS

In order to understand how CCNSW can improve its social media strategy on Twitter for the purpose of increasing its awareness and donations to reach the targets outlined in the previous section, it is important that we have a framework that enables us to understand the nature of twitter and its relation to charities. The following section will outline the R.O.P.E.S model which is a framework used by charities in the fundraising process. This section will explain and outline the important aspects of that model following the work of Kelly (1997) and Hougaard (2017) and then show how it can be applicable to a charity's social media strategy. This framework will enable us to devise a set of meaningful research questions for our analysis for the goal of improving CCNSW social media strategy on twitter.

R.O.P.E.S

The ROPES process is a normative and descriptive framework used for developing and implementing a fundraising process for charities (Kelly, 2001). Initially, Hougaard notes that this was developed by the New England Association of Catholic Development Officers. The acronym stands for Research, Objectives Programming, Evaluation and Stewardship. The figures (2) and (3) are a diagrammatic representation of the process and its workflow model.

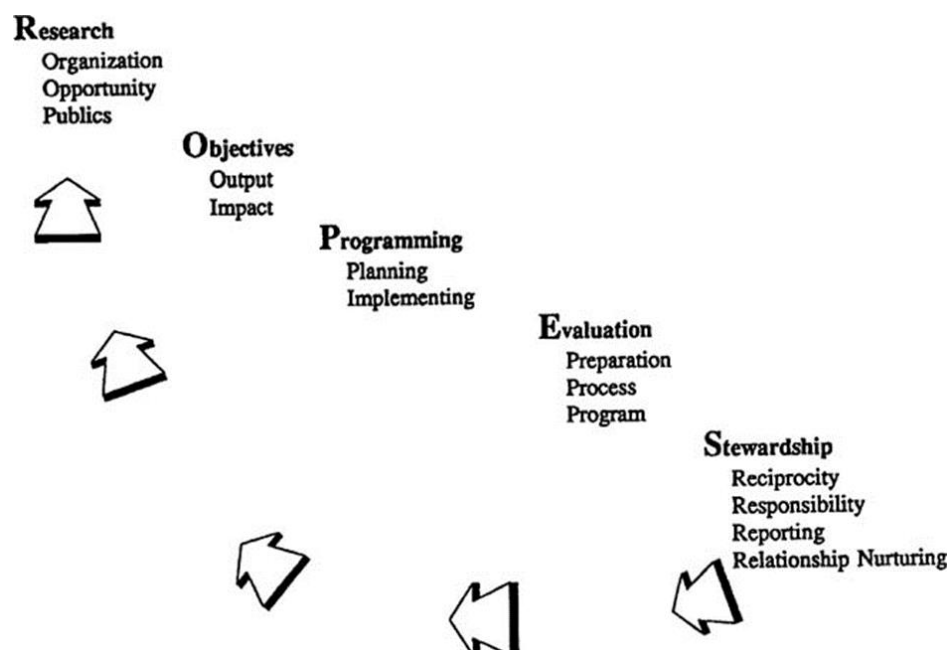


FIGURE 2: DIAGRAM OF THE R.O.P.E.S MODEL (KELLY, 2001)

As we can see from the figures, at each stage, there is a set of tasks our issues that need to be resolved before moving on to the next section. At the first level in the process there is research. Research is fundamental as without a solid understanding of the nature of the problem and understanding the dynamics of the situation it is difficult to achieve the results you want at the normative level (Kelly, 2001). In the case of charities, not having the correct data and knowledge of your supporters, services provided, and partners is problematic for developing sophisticated and accurate strategies later in the process (Kelly, 2001). This step is parceled into three different categories – organization, opportunity and publics.

Objectives is the second step after research. This is broken down into output and impact. These are related to the short- and long-term goals and impact the organization wishes to achieve, in the case of charities, what is the desired outcome given what is known from the research step? (Kelly, 2001) In the fourth step there is evaluation which is concerned with evaluating the outcomes from programming and its implementation. However, these three aspects of the ROPES process will not be part of the scope of this analysis.

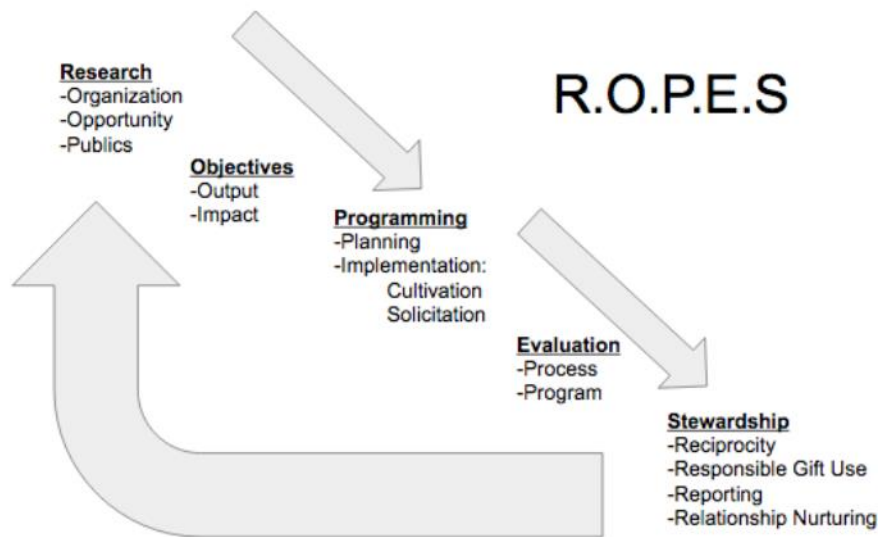


FIGURE 3: ANOTHER DIAGRAM WITH CULTIVATION AND SOLICITATION EXPLICIT IN THE IMPLEMENTATION PHASE OF PROGRAMMING. SOURCE: HOUGAARD (2017)

The two main elements of the ROPES processes framework that this report will utilize are Stewardship and Programming. These two elements are most readily seen in microblogging activities for charities and as such is helpful when developing the research questions.

In Programming there are two important elements for the fundraising process, these are part of the implementation subsection, these are cultivation and solicitation. (Hougaard, 2017, p 11) Cultivation refers to the relationship nurturing process that occurs in fundraising. With the end goal of getting donations being the main objective, it is important that a relationship is cultivated with the potential donors first. A charity relevant definition for cultivation can be defined as the process of encouraging the interest of a potential donor or volunteer using information about the charities institution and the services it provides (CFRE International, 2021, p 6)

Solicitation is related to the ‘ask’ or the act of asking from the charity. In many cases this is in the form of asking for donations or volunteering. This is arguably the most important step in the fundraising process, even though the other two fundraising aspects cultivation and stewardship take up more of the time. (Hougaard, p 14).

Lastly there is stewardship, this refers to the philosophy that an institution possesses regarding its ethical use of funds and resources donated to it from its donors and time from its volunteers. (CFRE International, p 14). A more practical definition that can be articulated in relation to charities is that stewardship refers to the act of thanking its donors and volunteers (Kelly, 2001). This act of thanking manifests itself in the explicit act of thanking (giving a shout-out to a set of volunteers or donors) but also in its transparency related to how the funds or volunteers are being utilized by the charity. This is stewardship as it helps to maintain the relationship with donors and volunteers by enabling them to see where their money or time is being utilized.

In twitter and microblogging in general, it is these three elements from out of the ROPES process which are most readily seen on social media. For example, the types of twitter posts that a charity can produce fall into one of these three categories or a combination of these categories exhaustively, meaning it is only possible for a charity to have posts related to its work that do one of three things – they cultivate relationships with prospective donors (cultivation), that ask for donations (solicitation) or that thank the donors (stewardship).

Given this theoretical understanding of the ROPES process we can now illuminate some research questions which can be borne out of this conceptual model as it relates to improving the CCNSW social media strategy on twitter to improve its overall donations and awareness.

RESEARCH QUESTIONS

The main research question(s) that will be part of the analysis is the following:

Q1 a): What is the percentage of engagement post types (cultivation, solicitation and stewardship) for Cancer Council NSW and the other charities overall?

Q1 b) What is the distribution of the twitter post likes? Mean number of likes per charity?

Q2: Which engagement type yielded the most likes? Which had the most retweets? Is this statistically significant?

Q3: Do posts with images improve likes and retweets?

Q4: Do posts with hashtags increase likes?

Questions 1 and 2 were formulated out of the ROPES framework. Question 2 is the most important research question in the report and is the main focus of the analysis. Questions 3 and 4 are also important but more general questions which would help in improving any of the solicitation, cultivation or stewardship type tweets for the CCNSW. From these questions, especially Q2-4 after analysis we will be able to provide recommendations to improve the social media strategy on twitter for CCNSW.

METHODOLOGY and ANALYSIS

The purpose of this methodology and analysis section was to answer the 4 questions that arose from the previous section. Question 2 is the most important research question and along with question 1, these are the types of questions that would be asked by a set of fundraisers who are following the ROPES process model and applying it to improving a social media campaign for a particular charity. Before we discuss the methodology of the analysis there are a certain set of assumptions that need to be considered that to help define the scope of the analysis.

ASSUMPTIONS

Assumption 1: The first assumption is that awareness will be operationalized to mean 'likes' or 'favorites'. In Twitter, there is a heart icon that allows somebody to 'like' or 'favorite' a particular post (see appendix A). The more Twitter users that 'like' or 'favorite' a post, the total number of likes will increase for that post. In this analysis, we will assume that liking a tweet is a metric to measure awareness.

The number of retweets a post has will also be secondary measure of awareness times only to compare it with the 'likes' or 'favorites' definition. This will only be used in some research questions to compare the differences if one uses retweets as the metric for determining awareness, however 'likes' as awareness will be the primary operationalization in this report unless stated otherwise.

The justification for this is that liking a tweet requires much less commitment compared to commenting and retweeting as it does not fill up your Twitter wall and therefore it is more likely people who retweet would also like the post, but not the converse, so the numbers of those aware of a post would be more accurately reflected in 'awareness as 'likes'' definition (Hellenkemper, 2017)

Assumption 2: Awareness correlates with donations. There is a correlation between the amount of total donations an organization receives from the public and the amount of awareness of the organization or charity (see Appendix B) (Caneva, 2011). If we take this as an assumption during our analysis, for the recommendation section, whatever strategies or tactics we will employ to improve awareness should also be considered something that will also benefit long term donation revenue as we will assume this correlation from the study.

METHODOLOGY

To answer the research questions posed above, an analysis of tweets from various charities was done using various statistical methods which will be mentioned later in this section of the report.²

The tweets sampled which were used in this analysis came from over 10 charities including CCNSW. These charities were: Guide Dogs Australia, Diabetes NSW & ACT, The Heart Foundation, Hepatitis Victoria (LiverWELL), The Movember Foundation, St John Ambulance (NSW), St Vincent de Paul (NSW), The Leprosy Mission Australia and the Deaf Society of NSW.

The purpose of having multiple charities was so that the generalization or insights derived from the data could be more universal and thus have greater applicability. By finding any trend or statistically significant result from the research questions over multiple charities, the inference would be stronger and therefore improve the strength of the recommendations.

These charities were selected based on their financial size. The Australian Charities and Not-for-profits Commission divides charities into three different subcategories by size – small, medium and large. Large charities are those with annual revenues of \$1 million dollars or more (Australian Charities Report, 2018, p 3). The total revenue in 2020 for Cancer Council Australia was \$21.7 million therefore it is considered a 'large' charity (Cancer Council Australia | Full Financial Report, 2020, p 12). As a result, its main competitors would be

² In the academic literature, this type of analysis is usually termed 'content analysis' see Bengtsson (2016) for more information.

those in the 'large' charity category and those charities are the ones that the Cancer Council NSW will be competing against for donors. All the charities chosen reached were part of the 'large charity' category of having an annual revenue of over \$1 million dollars AUD and were also in relatively similar competition spaces compared to CCNSW. In the appendix (Appendix C) there is a table showing annual revenue for 2018.

SAMPLES, DATA COLLECTION and ANALYSIS

The programming language R was used for the content analysis on the twitter posts. The packages 'rtweet', 'dplyr', 'stringr', 'base64enc', 'httpuv', 'gridextra', 'matrix' and 'tidyr' were all used as part of the web scraping, data cleaning and analysis process.

861 tweets were taken combined from the 10 charities twitter accounts. These 861 samples comprised of the working data set from the analysis would be undertaken to answer the research questions. The samples extracted from each twitter were from no earlier than February 2020 if they had under 100 tweets up to the present date or their last 100 tweets up to the present date if they had more. This was to ensure that the samples had some consistency across time. It was during February of 2020 that the covid pandemic became problematic in Australia (Parliament House Prime Minister | Press Release, 2020). Therefore, all the posts during that time to now would be formed under similar conditions. For example, as reported earlier in the SVA report, social distancing has affected the ability of charities to create and run events and meet ups and thus perhaps tweets from before the pandemic might have greater likes or may be of a different type (there may be more stewardship tweets when there is not a pandemic because there are more events thus more volunteers and donors that a charity can potentially thank). Only extracting tweets from this time period enabled the possibility of more accurate results that are relevant to recommendations in the present.

All posts that were retweets were not included in the twitter samples. This was because it was important to get tweets that were uniquely constructed by the charities themselves, not those by others which may have been irrelevant to the ROPES model of categorizing tweets under the cultivation, stewardship and solicitation umbrella. Additionally, a charity may retweet a topic that is not relevant to the charity or organization and the likes on the post may not represent awareness of the charity and its mission but rather a different issue altogether. For these reasons, retweets of other twitter accounts posts were excluded from the samples.

Once the 861 tweets (samples) were gathered from the 10 charities the process of categorizing the tweets into stewardship, cultivation or solicitation tweets was done manually. There are advanced learning machine techniques that could be utilized to automate this process however developing an algorithm was time consuming and the accuracy was paramount to this task, therefore a manual approach was more optimal for this analysis.³ An objective metric that evaluated the content of each post or tweet was used to determine which of the three categories it would fall under. Borrowing the metric developed by Hougaard (2017) was used to do this, which can be seen in the appendix (Appendix D). However, unlike the Hougaard study, there was no mixed categories – a decision was made in this report to cleanly divide each post into one of the three categories depending on how many points it scored on the objective metric sheet for that category. Which category it had the most points in (e.g. if it had more solicitation elements than stewardship elements it was categorized as a solicitation tweet).

The posts once categorized into the three separate categories was then put into a dataframe on R with the rest of the metadata that was available to be extracted using the 'rtweet' package. The first research question was of a more exploratory nature and as such a rectangular tree graph was used to represent the data for its second part.⁴

For research question 2 the posts were categorized into their respective tweet types as defined by the ROPES process (cultivation, stewardship and solicitation) alongside the posts metadata. The mean number of likes for each per post for each category was then computed and a one-way ANOVA test was used to determine

³ One such approach is using the LDA algorithm see Kelechava (2019).

⁴ For why rectangular tree maps are more effective than pie charts, see Healy (2018).

statistically significance. Tukey's HSD was then used to determine which of the three categories had a statistically significant difference.

For research questions 3 and 4 the data was re-transformed and split the two respective groups into those with images vs no images and then posts with hashtags and those without hashtags. An unpaired t-test was used for each question to determine if there was a significant difference among groups in terms of mean number of likes.

RESULTS

QUESTION 1

PART A

The table 1 shows the distribution pattern of the types of twitter posts from the sample. Out of the 861 samples used in the analysis 134 fell under being a solicitation type tweet after categorization, which accounts for 15.6% of total tweets. For stewardship type posts, 252 or 29.3% were categorized under that label. And lastly, for cultivation posts 475 or 55.2% of the samples fell into this category after the manual analysis that was undertaken with help from the metric developed by Hougaard (2017).

Post Type	Number of Posts	Percentage of Total
Solicitation	134	15.60%
Stewardship	252	29.30%
Cultivation	475	55.20%

TABLE 1: DISTRIBUTION OF POSTS VIA TWEET TYPE FROM OVERALL SAMPLE (N = 861)

Post Type	Number of Posts	Percentage of Total
Solicitation	17	32.30%
Stewardship	14	26.90%
Cultivation	21	40.10%

TABLE 2: DISTRIBUTION OF POSTS VIA TWEET TYPE FOR CCNSW (N = 52)

In comparing the overall distribution of tweet types from the sample compared to the CCNSW specifically we can see differences. Table 2 shows the results for CCNSW tweets. There were 52 samples from the CCNSW. Solicitation posts accounted for 17 posts or 32.3%. Stewardship was 14 or 26.9% and cultivation was 21 or 40.1% of the total tweets by the CCNSW.

In terms of total proportion, both CCNSW compared with the samples overall both had cultivation posts as the majority in both cases. However, the stewardship and solicitation posts were switched. For the overall distribution stewardship posts were ranked in second and solicitation were ranked last, for CCNSW specific posts, stewardship posts were ranked last while solicitation posts were ranked second.

PART B

Part two of question 1 investigated the distribution of the likes and the mean number of likes per charity. The table and histogram below show visually the results. As can be seen (figure 4) the distribution pattern indicates a positive skew. This means that the majority values are at the lower end with the higher values reducing the further along the x axis. Very few twitter posts have values at the higher end of the likes count.

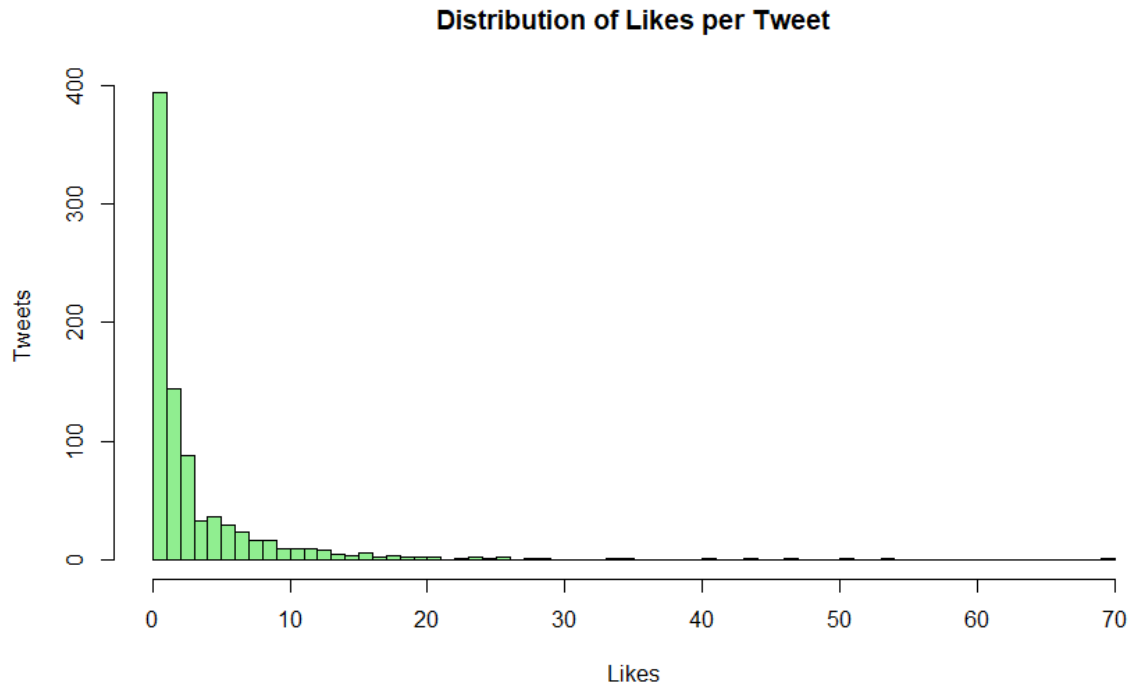


FIGURE 4: DISTRIBUTION OF LIKES PER TWEET FROM THE SAMPLE (N = 861) THIS HISTOGRAM PLOT SHOWS THE POSITIVELY SKEWED DISTRIBUTION.

From table 3 we can see the distribution of mean likes per twitter post from each charity. The Movember Foundation had the highest number of likes per tweet with a mean of 9.59 while the Deaf Society NSW had the lowest number of likes per tweet at 0.56.

Charity	Mean Likes per Post	Percentage of Total
VinnesNSW	1.93	5.74%
Leporsy Mission	1.09	3.24%
St Johns Ambulance NSW	3.17	9.43%
Movember	9.59	28.52%
LIVERwell	1.44	4.28%
Heart Foundation	9.05	26.92%
Diabetes NSW	0.94	2.80%
Deaf NSW	0.56	1.67%
CCNSW	3.69	10.98%
Guide Dogs Aus	2.16	6.42%

TABLE 3: AVERAGE AMOUNT OF LIKES FOR A POST FOR EACH CHARITY.

A treemap (see figure 5) was also generated to illustrate the mean number of likes per twitter post with the physical size of the tiles representing the quantity or number of likes per tweet each charity received.

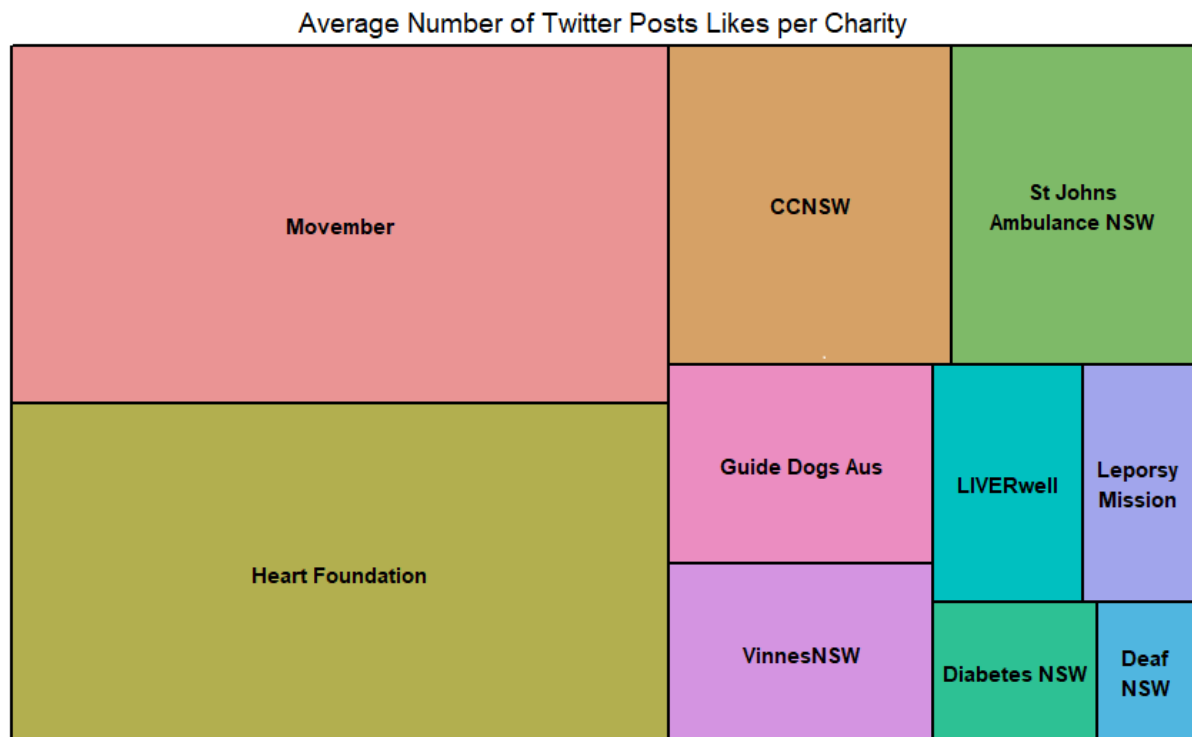


FIGURE 5: A TREEMAP OF THE AVERAGE NUMBER OF TWITTER POSTS LIKES FOR EACH CHARITY .

QUESTION 2

The number of likes per tweet and the standard deviation for each tweet category is shown in table 4. Cultivation tweet posts generated an average of 3.3 likes, stewardship posts 4.6 and solicitation tweets 2.9 likes per post.

Tweet Type	Mean Likes per Post	SD of Likes per Post
Cultivation	3.3	5.61
Stewardship	4.6	7.89
Solicitation	2.9	3.87

TABLE 4: MEAN AND STANDARD DEVIATION OF LIKES FOR EACH TWEET CATEGORY .

From inspection stewardship type tweets generated the most likes at 4.6 per post. However, in order to make sure the result was statistically significant a one-way ANOVA was performed on the data to test whether the difference in the means between the three groups were statistically significant.

Figure 6 was created to see whether the normality assumption held for the residuals of the data. As one can visually see, the distribution of residuals tends to a normal distribution except with a fat tail on one side, however it is reasonable to believe that the normality assumption is not violated in this case. The QQplot was generated to see visually to see if the data looked normal. A boxplot (figure 7) was also created to show the distribution of the groups and the number of likes per post.

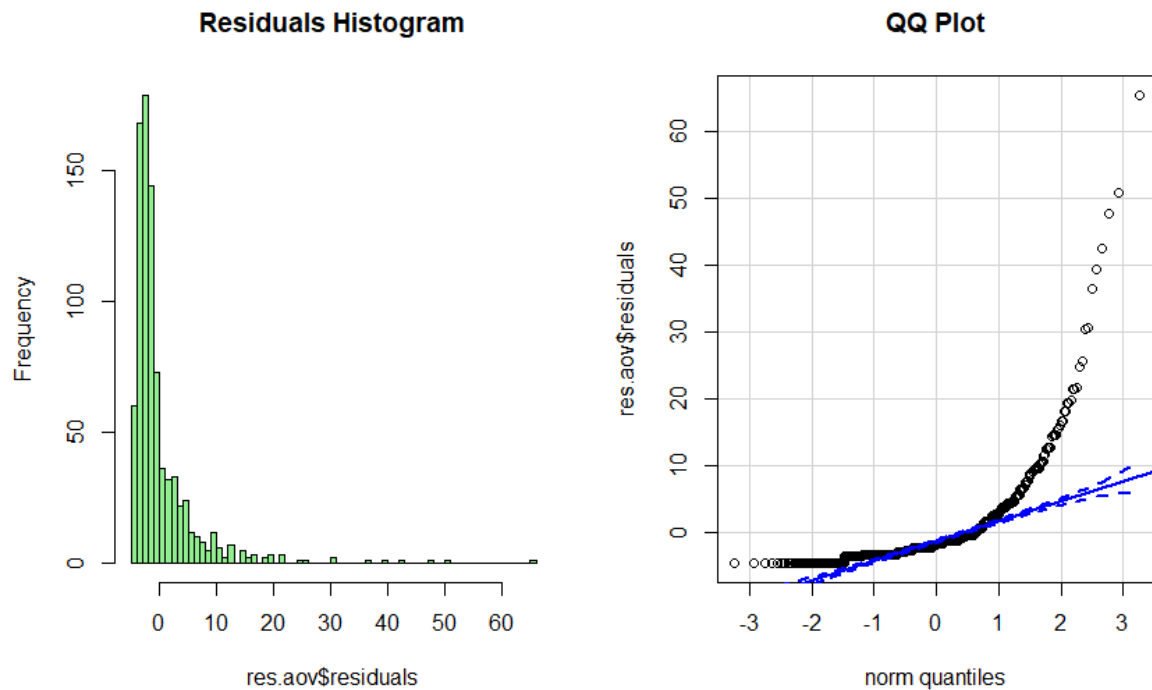


FIGURE 6: PLOT 1 (LEFTSIDE) IS A RESIDUALS HISTOGRAM, PLOT 2 (RIGHT SIDE) IS A QQ PLOT OFTEN USED TO TEST THE NORMALITY OF THE DATA.

The results from the one-way ANOVA test yielded a p-value of less than 0.001. Following the standard cut off of 0.05, the results show that there is a statistically significant difference between the three different group means (the groups being the solicitation, stewardship and cultivation type tweets).

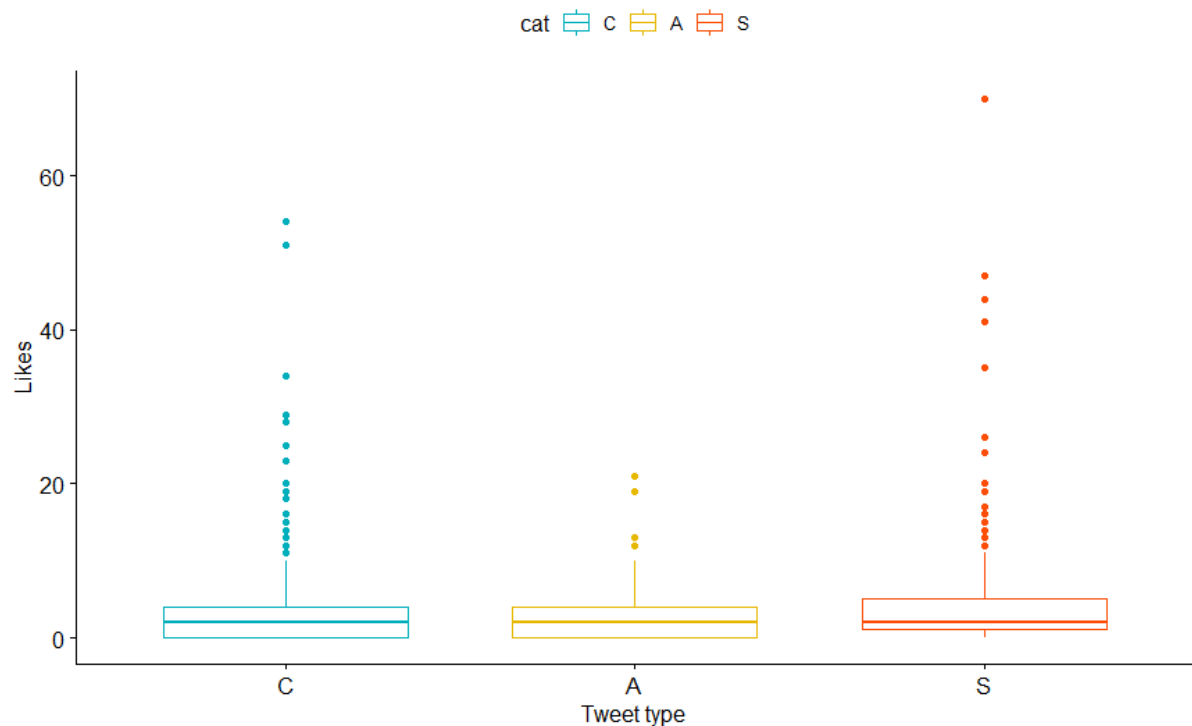


FIGURE 7: MULTIPLE BOXPLOTS REPRESENTING EACH TWEET CATEGORY AND THE DISTRIBUTION OF THEIR DATA.

However, in order to determine which groups specifically were statistically significant in the difference of their means, Tukey's honest significance test was also performed. For stewardship vs cultivation posts the

difference in likes between the two groups was given a p value of 0.024. for stewardship vs solicitation posts, the difference in means between the groups was given a p value of 0.028. For cultivation vs solicitation tweets a p value of 0.75 was generated. The results showed that the statistically significant difference is between cultivation and stewardship and stewardship and solicitation. No significant difference between solicitation and cultivation tweets. Therefore, the results suggest that stewardship tweets get more likes compared to cultivation and solicitation tweets.

In the interest of completion, a one-way ANOVA was also performed to see whether there was a statistically significant difference between the mean number of retweets among the three groups. A p-value of 0.77 was calculated, therefore it was not a statistically significant result.

QUESTION 3

The third question divided up the 861 tweets into two groups, those that had images and those that did not. This was identified using a function from 'rtweet' which used the metadata to determine if there was images or video in the twitter post. Once the two groups were created, the table 4 shows the mean number of likes per twitter post by group. For posts with images, the average amount of likes was 4.23 and for posts without images the average amount of likes was 2.89.

Tweet Media Type	Mean Likes per Post
Images	4.23
No Images	2.89

TABLE 5: MEDIA TYPE AND THE AVERAGE LIKES PER POST

An unpaired t-test was then performed to test for statistical significance. A p value of 0.001 was calculated meaning that the difference in means were statistically significant assuming a standard cut off value of 0.05. This suggests that the twitter posts with images generate more likes on average than those without images, irrespective of the twitter type they are (cultivation, solicitation or stewardship).

QUESTION 4

The last research question looked at whether there was a difference in the number of likes generated with twitter posts with hashtags vs ones without. The dataframe was once again categorized under those posts which had hashtags and ones that did not, irrespective of their twitter category type. The mean number of likes for twitter posts with hashtags was 3.8 while without hashtags was 3.4 which can be seen in table 6.

Hashtags	Mean likes per Post
No	3.42
Yes	3.79

TABLE 6: AVERAGE NUMBER OF LIKES PER POST FOR TWEETS WITH HASHTAGS VS NO HASHTAGS.

An unpaired t-test was performed on the data. The calculated p value was 0.38. Therefore, this was above the outside of the cut off range for being statistically significant (where $p < 0.05$). This suggests that there is no difference between twitter posts with hashtags versus those without.

DISCUSSION and RECOMMENDATIONS

A statistical analysis was performed on the data and the results section generated some interesting findings. Firstly, two of the four research questions were developed using the ROPES framework as the basis for formulating the questions. Following Hougaard (2017), the idea of categorizing tweets into cultivation stewardship and solicitation tweets only made sense under the context of ROPES. Applying some of the conceptual ideas from ROPES enabled us to develop questions which would be useful in generating insight from the data which pertained to fundraising and awareness.

The second two questions were more ancillary questions which would nevertheless help the CCNSW in improving their general social media strategy on twitter for the purposes of fundraising and generated awareness.

The first question researched the distribution of tweet types overall compared to CCNSW individually. From the data it shows that many tweets (CCNSW individually and all the charities collectively) were cultivation tweets. The purpose of cultivation tweets is to get potential donors and volunteers to join the cause of the charity. This is done through information dissemination about the charity and its mission or aims (cite, x5 and friends with money). This suggest that charities are using their twitter primarily for the purposes of getting new followers and people interested in their cause. This makes intuitive sense, since social media is used as form of marketing for both non-profit and profit organizations and marketing is always about increasing its customer base or in the case of charities, increasing its donor and volunteering base (McGrath, 2016). However, CCNSW did have much more solicitation type tweets compared to the general distribution. This could be for a multitude of reasons – perhaps CCNSW is much more comfortable with asking for donations or it could be a result of sampling issues.

The second part of the first question looked at the number of likes per tweet per organization. The Movember Foundation had the most likes per tweet with CCNSW coming in at third highest. Looking at how the Movember Foundation was able to generate almost 3 times the amount of likes per tweet could be something that CCNSW could investigate in the future to replicate the success the Movember Foundation has had in generated so many likes. One issue however is that there could be outliers in the data which are affecting the general trend. If it is the case that one or two tweets had exponentially larger amounts of likes compared to the average for the Movember foundation, this large average could be a result of special cases, rather than Movember implementing successful social media tactics in their posts that CCNSW would want to emulate and replicate.

The second question looked at what type of tweet post (cultivation, stewardship or solicitation) yielded the most likes. This question along with question 1 was directly inspired by using the ROPES framework which is a way for charities to structure the ontology of fundraising. By viewing fundraising through the ontology provided by the ROPES framework, previous work (from Hougaard 2017) realized that the stewardship and cultivation and solicitation ideas within the ROPES framework was manifested in social media through the types of tweets that charities could theoretically post. Therefore, doing a statistical analysis of what tweet type generated the most likes would be insightful for CCNSW for the purposes of increasing awareness and donations to offset the loss in donation income from the pandemic. Unlike the Hougaard study, there was a decision to only have 3 categories and any subcategories. The reasons for this in the report were that it would be unfeasible and impractical to implement any recommendations for the CCNSW if the model were too complex.

It was found in the results that stewardship tweets generated the most likes compared to solicitation and cultivation tweets. One could reason that this was because it is more likely a person would be interested in liking a tweet when they see the organization giving thanks to a supporter, volunteer or donor. As these types of people would be viewing the twitter accounts, they would also find satisfaction in being thanked. Retweets were also investigated but they did not give any statistically significant results. This meant that posts from the

CCNSW which were retweeted versus ones that were didn't receive on average a difference in the number of likes.

The last two questions were secondary research questions which were not necessarily borne out of the ROPES framework but nevertheless they were interesting questions which could help CCNSW. The first question looked at whether images improved the number of likes on a post. It was found that to be case that it did. This tends to align with current research, people tend to be drawn and recall better media which includes images other than just plain text (Hockley, 2008).

The last question looked at whether hashtags improved the amount of likes a twitter post generated. The results from the statistical analysis showed that there was no difference in hashtag posts versus those without hashtags. This result was surprising. Literature on the social media marketing, quite often make the claim that hashtags should improve a twitter post or persuade people to donate (See Soboleva et al, 2017, p 521). But the results from this report show that this does not translate to any objective measure in terms of the amount of likes a post receives. More investigation into this line of research would be fruitful as there could be sampling issues involved or perhaps the dynamics of how hashtags improve twitter posts is more complex than just the amount of likes a post receives.

From this, we can develop a set of recommendations for the CCNSW after the analysis and results in the report. Firstly, some justification needs to be done in what courses of action should be taken in priority given all the recommendations cannot be implemented.

The number of recommendations given will be four. These represent recommendations from insight gained from each of the four research questions. The course of action which will be highest should is firstly dictated by if the research question was generated by the ROPES model AND if there was statistical test. Question 2 was positive on both criteria; it was a question which was formulated under the ontology of the ROPES framework, and it also had a statistical test. The second highest course of action is one that had a statistical test OR was generated from the ROPES framework. Both questions 1 and 3 fit this criterion. However, since question 3 had a statistically significant result, that will take priority over any recommendation developed from question 1. Lastly, insight for recommendations developed from question 4 will be given lowest priority as it was not generated by the ROPES model, nor did it have a statistically significant result.

RECOMMENDATION	Course of Action
HIGHEST	Generate more stewardship tweets.
HIGH	Add images to twitter posts
LOWER	Emulate tweeting style of Movember Foundation
LOWEST	Add hashtags to tweets

TABLE 7: SUMMARY OF RECOMMENDATIONS AND THEIR PRIORITY LEVELS

The table 7 summarizes the courses of action. Firstly, CCNSW should generate more stewardship tweets. This is important as the data suggests that these yield the most likes. This means creating more stewardship tweets on an absolute scale, whilst also looking into the qualitative aspects of stewardship posts to improve their overall reception. By doing this, they should be able to generate more likes and therefore more awareness and this should translate into more donations in the long term, given our assumptions made previously in the report about how awareness correlates with total donations. Each financial year or cycle, there should be a review on if stewardship tweets are generating more likes as times goes on, both on a relative and absolute measure. Over the process of 3-4 years, they should develop metrics to see if this social media course of action is helping them improve their long-term donation revenue, and in the short term if it is helping them goal of mitigating the projected 30% drop in donation funds within the current COVID-19 pandemic.

The second course of action they should take is to add images to their twitter posts. While they already add images to their twitter posts, by making sure that every twitter post has an image, this may guarantee greater likes, irrespective of the type of tweet that it is. For further lines of research that may help, CCNSW should investigate whether particular types of images generate more likes than other types of images, by getting into this they will be able to be more efficient and effective in their image curation for tweets. Again, the purpose is to adding images to posts is to generate more likes which is a metric of awareness and therefore will hopefully lead to further donations in the long run.

The third recommendation is that they emulate the tweeting style or see if they can find elements within the Movember Foundations twitter page that makes them so successful in generating likes. By taking important elements that they can identify that can generate likes, they can then use these findings and customize them to their own goals and needs for the purposes of raising awareness and donations.

Lastly, CCNSW should adopt the use of hashtags in their twitter posts if they feel the need though this is lowest on the priority scale. While hashtags seeming to be relatively unimportant is counter-intuitive and literature suggests the opposite.⁵ The fact that this analysis did not find any statistically significant difference in tweets which had hashtags compared to those that did not regarding average likes, suggests that this may be an unimportant feature statistically speaking. However further research can be done into this to make sure that this is the case. One study does not determine the truth or consensus, but rather the meta-analysis of the body of literature on that topic should. However, taking this report on its merits and its results seriously, the results indicates that hashtags are not as important as previously suggested and therefore should be lower on the priority scale in their course of action of improving their social media strategy.

In conclusion, the aim and goal of this report was to improve CCNSW's social media strategy on twitter to increase awareness and donations. Since the report was written during the COVID-19 pandemic which is still currently happening at the time of this report, the goal of their social media strategy was to mitigate the projected 30% decrease in fundraising revenue over the next year. The ROPES framework which is used by charities to develop a fundraising model and/or campaign gave us an ontology to view how twitter functions for charities. Three important concepts from the ROPES framework and process - cultivation, stewardship and solicitation were seen to be manifested in social media in the form of tweets by charities. Four research questions were then developed, and a statistical analysis was performed. The recommendations were then provided in the form of a hierarchy outlined above.

⁵ See Social Misfits Media (2014, p 3).

REFERENCES

Social Ventures Australia and the Centre for Social Impact, 2020. *Will Australian charities be COVID-19 casualties or partners in recovery? A financial health check*. [online] UNSW. Available at: <https://www.csi.edu.au/media/Charities_and_Covid-19_Report.pdf> [Accessed 23 April 2021].

Bengtsson, M., 2016. How to plan and perform a qualitative study using content analysis. *NursingPlus Open*, 2, pp.8-14.

Cancer Council NSW, 2020. *Financial Report 2019/2020*. [online] Cancer Council NSW. Available at: <<https://www.cancercouncil.com.au/wp-content/uploads/2021/02/CCNSW-2019-20-Financial-Report-final.pdf>> [Accessed 1 June 2021].

Cancer Council Australia, 2020. *Annual Review 2019/2020*. [online] Cancer Council Australia. Available at: <https://www.cancer.org.au/assets/pdf/cca-annual-report-2019-2020#_ga=2.41965408.453237344.1623840543-79486430.1621692387> [Accessed 19 May 2021].

Cancer Council Australia, 2021. *Our history*. [online] Available at: <<https://www.cancer.org.au/about-us/about-cancer-council/australia/history>> [Accessed 27 April 2021].

Caneva, L., 2021. *Building Awareness Key to Donation Success - Report | PBA*. [online] Pro Bono Australia. Available at: <<https://probonoaustralia.com.au/news/2015/02/building-awareness-key-to-donation-success-report/>> [Accessed 28 April 2021].

Cfre.org. 2021. *Glossary of Fundraising Terms*. [online] Available at: <<https://www.cfre.org/wp-content/uploads/2018/10/Handout-Glossary-of-Fundraising-Terms.pdf>> [Accessed 10 June 2021].

Changepath.com.au. 2021. *Cancer Council NSW - organisation details*. [online] Available at: <<https://www.changepath.com.au/cancer-council-nsw>> [Accessed 30 April 2021].

Healy, Y., 2018. *The issue with pie chart*. [online] Data-to-viz.com. Available at: <<https://www.data-to-viz.com/caveat/pie.html>> [Accessed 4 June 2021].

Hellenkemper, M., 2017. *Ignoring Engagement On Social Networks Could Be Disastrous For Your Influencer Marketing Campaign!.* [online] Business 2 Community. Available at: <<https://www.business2community.com/instagram/ignoring-engagement-social-networks-disastrous-influencer-marketing-campaign-01852513>> [Accessed 5 June 2021].

Hockley, W.E., 2008. The picture superiority effect in associative recognition. *Memory & cognition*, 36(7), pp.1351-1359.

Hougaard, S., 2017. *Tweeting for a Cause: A Content Analysis of Successful Charitable Nonprofits' Publishing Strategies on Twitter.* [online] Brigham Young University. Available at: <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=7274&context=etd> [Accessed 15 May 2021].

Johns, G., 2020. *Australian Charities Report Shows the Importance of the Sector.* [online] Australian Charities and Not-for-profits Commission. Available at: <<https://www.acnc.gov.au/media/news/australian-charities-report-shows-importance-sector>> [Accessed 4 June 2021].

Kelechava, M., 2019. *Using LDA Topic Models as a Classification Model Input.* [online] Medium. Available at: <<https://towardsdatascience.com/unsupervised-nlp-topic-models-as-a-supervised-learning-input-cf8ee9e5cf28>> [Accessed 6 June 2021].

Kelly, K.S. 1997. *Effective Fund-Raising Management* (1st ed.). Routledge. <https://doi.org/10.4324/9781410602664>

Kelly, K. 2001. Stewardship: The Fifth Step in the Public Relations Process. In: 2001. *Handbook of Public Relations*, Thousand Oaks, CA: SAGE Publications, Inc. pp. 279-290 Available at: <<http://www.doi.org.ezproxy.uws.edu.au/10.4135/9781452220727.n21>> [Accessed 16 Jun 2021].

McGrath, J., 2021. *How to Use Social Media Marketing to Grow Your Customer Base.* [online] Blog.marketveep.com. Available at: <<https://blog.marketveep.com/how-to-use-social-media-marketing-to-grow-customer-base>> [Accessed 9 June 2021].

Prime Minister of Australia, 2020. *Transcript Press Release*. [online] Available at: <<https://www.pm.gov.au/media/press-conference-australian-parliament-house-4>> [Accessed 6 June 2021].

Soboleva, A., Burton, S., Daellenbach, K. and Basil, D.Z., 2017. Tweets for tots: Using Twitter to promote a charity and its supporters. *Journal of Consumer Marketing*.

Svensson, P. G., Mahoney, T. Q., & Hambrick, M. E., 2014. *Twitter as a Communication Tool for Nonprofits*. *Nonprofit and Voluntary Sector Quarterly*, 44(6), 1086–1106. doi:10.1177/0899764014553639

Social Misfits Media, 2014. *Friends with Money: A guide to fundraising on social media*. 1st ed. [Online] Social Misfits Media. Available at: <[https://hubble-live-assets.s3.amazonaws.com/cranfield-trust/attachment /file/18/Friends_With_Money__2_. pdf](https://hubble-live-assets.s3.amazonaws.com/cranfield-trust/attachment/file/18/Friends_With_Money__2_.pdf)> [Accessed 10 June 2021]

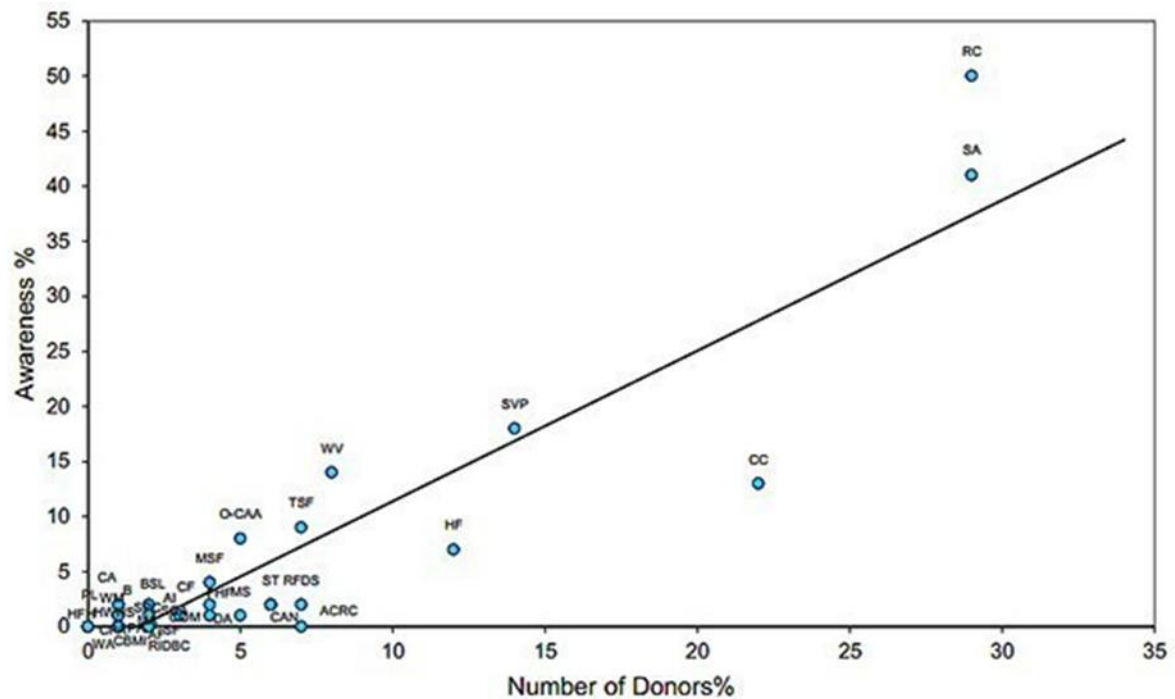
APPENDIXES

APPENDIX A. The “likes” button and count on a Twitter post.



SOURCE: CANCER COUNCIL NSW, 2021

APPENDIX B. Graph showing the correlation between donations and awareness.



Key:

AJ	Amnesty International	MS	MS Society
AF	Angel Flight	O-CAA	Oxfam/Community Aid Abroad
ACRC	Australian Cancer Research Council	PA	Parkinson's Australia
B	Barnardos	PL	Plan - Child Sponsorship/Plan International
CC	Cancer Council	RC	Red Cross
CAN	CanTeen	RFDS	Royal Flying Doctor Service
CA	Care Australia	RIDBC	Royal Institute for Deaf & Blind Children
CF	Careflight	SA	Salvation Army
CPA	Cerebral Palsy Alliance	SCA	Save the Children Australia
CBMI	Christian Blind Mission International	SPC	Spastic Centre
COM	Compassion Australia	SVP	St. Vincent de Paul/Vinnies
DA	Diabetes Australia	ST	Starlight Foundation
FHF	Fred Hollows Foundation	SF	Stroke Foundation
HFH	Habitat for Humanity	BSL	The Brotherhood of St Laurence
HF	Heart Foundation	TSF	The Smith Family
HWNS	House with no Steps	WA	Water Aid Australia
MSF	Medecins Sans Frontieres/MSF Australia/Doctors Without Borders	WM	Wesley Mission
MA	Mission Australia	WV	World Vision

SOURCE: CANVEA (2021)

APPENDIX C. Annual revenue for charities in 2018

Charity	Total Revenue (2018)
VinnesNSW	\$204,063,383.00
Leporsy Mission	\$5,435,476.00
St Johns Ambulance NSW	\$22,382,000 ²
Movember	\$31,412,221 ²
LIVERwell	\$2,308,695 ²
Heart Foundation	\$74,981,000.00
Diabetes NSW	\$17,358,415.00
Deaf NSW	\$7,475,601.00
CCNSW	\$73,378,000.00
Guide Dogs Aus	\$1,624,546 ²

VALUES ARE FROM SOURCE: CHANGEPATH.COM.AU

APPENDIX D. Criteria for determining tweet category type by Hougaard.

Cultivation	No	Yes
Disclosure/Info		
Dissemination	No	Yes
Purpose/Mission	No	Yes
Link to Site/Social Pages	No	Yes
Link to Media Coverage	No	Yes
Link to Announcements	No	Yes
Illustration/Video of Info	No	Yes
Involvement (encouraged)	No	Yes
Solicitation	No	Yes
Money	No	Yes
Time	No	Yes
Service	No	Yes
Get Involved	No	Yes
Stewardship	No	Yes
Reciprocity	No	Yes
Responsibility	No	Yes
Reporting	No	Yes
Relationship Nurturing	No	Yes

SOURCE: (HOUGAARD 2017)

APPENDIX E. Analysis Results from R.

```
> res.aov
Call:
aov(formula = favorite_count ~ cat, data = massive_df)

Terms:
          cat Residuals
Sum of Squares   350.22 32906.16
Deg. of Freedom      2     858

Residual standard error: 6.192913
Estimated effects may be unbalanced
```

```
> TukeyHSD(res.aov)
Tukey multiple comparisons of means
 95% family-wise confidence level

Fit: aov(formula = favorite_count ~ cat, data = massive_df)

$cat
      diff      lwr      upr      p adj
C-A 0.433857 -0.9883481 1.856062 0.7539628
S-A 1.700190  0.1456786 3.254700 0.0280375
S-C 1.266332  0.1332203 2.399445 0.0240103
```

```
> t.test(photo_group$`massive_df$favorite_count`, no_photo_group$`massive_df$favorite_count`)

welch Two sample t-test

data: photo_group$`massive_df$favorite_count` and no_photo_group$`massive_df$favorite_count`
t = 3.1405, df = 772.61, p-value = 0.001751
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.5056079 2.1914858
sample estimates:
mean of x mean of y
 4.228866  2.880319
```

Welch Two Sample t-test

```
data: no_hashtag_likes and hashtag_likes
t = 0.87693, df = 827.23, p-value = 0.3808
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.4521162  1.1823374
sample estimates:
mean of x mean of y
 3.787524  3.422414
```

```
# MASTERS PROJECT CODE
```

```
# Twitter Key
```

```
key = "Wtn7cEgiT5okpz8qljQ1nbNzJ"
```

```
secret = "ariCoPDXMCov0SWacAjifw3zwsVE3tbkfmHkhJ5Dap7VHgSvto"
```

```
#API key:
```

```
#Wtn7cEgiT5okpz8qljQ1nbNzJ
```

```
#API secret key:
```

```
#ariCoPDXMCov0SWacAjifw3zwsVE3tbkfmHkhJ5Dap7VHgSvto
```

```
#=====
```

```
# AUTHORISING R #
```

```
#=====
```

```
library("rtweet")
```

```
library("dplyr")
```

```
library("stringr")
```

```
library("tm")
```

```
library("base64enc")
```

```
library("httpuv")
```

```
library("gridExtra")
```

```
library("Matrix")
```

```
library("tidyr")
```

```
library("treemap")
```

```
library("ggpubr")
```

```
if(!require(devtools)) install.packages("devtools")
```

```
devtools::install_github("kassambara/ggpubr")
```

```
install.packages("ggpubr")
```

#Set up OAuth

```
create_token(  
  app = "my_twitter_app",  
  consumer_key = key,  
  consumer_secret = secret)
```

#TWEETS, EARLIEST FEB 2020 or LAST ONE BEFORE FEB IF NO FEB TWEETS ELSE LAST 100 TWEETS

Cancer Council NSW

#"2020-02-04 03:19:03 UTC" From this beginning date to present

```
tl_ccnsw <- get_timelines("CCNewSouthWales", n = 1000, include_rts = FALSE)
```

```
tl_ccnsw$text[1:52]
```

Tweets from closest to present day back since 4th of February 2020.

```
tl_ccnsw_cd <- tl_ccnsw[1:52, ]
```

Deaf Society NSW

```
tl_dfsnsw <- get_timelines("deafsocietynsw", n = 1000, include_rts = FALSE)
```

Tweets from closest to present day back since 23rd September 2019

```
tl_dfsnsw_cd <- tl_dfsnsw[1:9, ]
```

St John Ambulance NSW

```
tl_sja <- get_timelines("stjohnnsw", n = 1000, include_rts = FALSE)
```

From present day to 8th of March 2021

```
tl_sja_cd <- tl_sja[1:100, ]
```

```
# Heart Foundation
```

```
tl_hf <- get_timelines("heartfoundation", n = 1000, include_rts = FALSE)
```

```
#4th of April 2021 first tweet to present
```

```
tl_hf_cd <- tl_hf[1:100, ]
```

```
# The Leporsy Mission
```

```
tl_tlma <- get_timelines("TLMAustralia", n = 1000, include_rts = FALSE)
```

```
#17th of december 2020 to present (last 100)
```

```
tl_tlma_cd <- tl_tlma[1:100, ]
```

```
#St Vinnies NSW
```

```
tl_vnsw <- get_timelines("VinniesNSW", n = 1000, include_rts = FALSE)
```

```
#24th of February 2021 to present (last 100)
```

```
tl_vnsw_cd <- tl_vnsw[1:100, ]
```

```
# Guide Dogs NSW
```

```
gd_nsw <- get_timelines("GuideDogsAus", n = 1000, include_rts = FALSE)
```

```
# 22nd of March 2021 to present (last 100)
```

```
gd_nsw_cd <- gd_nsw[1:100, ]
```

```
# Movember Foundation
```

```
tl_mo <- get_timelines("MovemberAus", n = 1000, include_rts = FALSE)
```

```
# November 12 2020 (last 100)
```

```
tl_mo_cd <- tl_mo[1:100, ]
```

```
#LIVERwell incorporating victoria hepatitis
```

```
tl_lw <- get_timelines("hepvic", n = 1000, include_rts = FALSE)
```

```
# 6th of May 2020 to present (last 100)
```

```
tl_lw_cd <- tl_lw[1:100, ]
```

```
# Diabetes NSW
```

```
tl_dnsw <- get_timelines("DNSW_ACT", n = 1000, includte_rts = FALSE)
```

```
#30th of march 2021 to present (last 100)
```

```
tl_dnsw_cd <- tl_dnsw[1:100, ]
```

```
#Tweet_type column
```

```
#Cancer Council tl_ccnsw_cd
```

```
#Cultivation - C, Solicitation - A (ask), Stewardship - S
```

```
cancer_council_cat <-
```

```
c("C", "S", "S", "S", "S", "C", "A", "S", "C", "A",
```

```
"S", "S", "S", "A", "C", "C", "C", "C", "C", "A",
```

```
"S", "A", "C", "C", "C", "A", "C", "C", "C", "C",
```

"C", "S", "A", "A", "C", "C", "A", "A", "A", "A",

"A", "S", "S", "A", "S", "A", "A", "C", "C", "S",

"A", "C") # 52 total.

Guide Dogs Australia tweet type

guide_dogs_cat <-

c("A", "A", "C", "C", "C", "C", "C", "C", "C", "S",

"S", "C", "S", "S", "C", "S", "S", "S", "C", "S",

"S", "A", "S", "C", "S", "C", "C", "C", "C", "C",

"S", "S", "S", "S", "S", "S", "S", "S", "S", "C",

"C", "C", "C", "C", "C", "S", "A", "C", "C", "S",

"C", "A", "C", "C", "C", "C", "C", "C", "C", "C",

"C", "C", "C", "C", "S", "S", "S", "C", "C", "C",

"C", "C", "C", "S", "S", "S", "S", "S", "C", "C",

"C", "C", "C", "C", "S", "C", "S", "S", "S", "C",

"C", "C", "C", "C", "C", "A", "A", "C", "C", "S")

diabetes_nsw_cat <-

c("C", "C", "C", "S", "A", "C", "C", "C", "C", "C",

"C", "C", "C", "S", "C", "C", "C", "C", "S", "C",

"S", "C", "C", "C", "C", "C", "C", "C", "C", "C",

"C", "C", "C", "C", "S", "S", "C", "C", "S", "C",

"C", "C", "C", "S", "C", "S", "C", "C", "C", "C",

"C", "C", "C", "S", "C", "C", "C", "A", "C", "C",

"C", "C", "C", "C", "C", "C", "S", "C", "C", "C",

"C", "C", "C", "C", "C", "C", "C", "C", "C", "C",

"C", "C", "C", "C", "C", "C", "C", "C", "C", "C",

"C", "C", "A", "C", "C", "C", "C", "C", "C", "C")

heart_foundation_cat <-

c("A", "A", "C", "C", "S", "C", "C", "C", "C", "A",

"C", "S", "S", "S", "C", "C", "S", "C", "C", "C",

"C", "C", "S", "C", "C", "S", "S", "C", "C", "C",

"C", "C", "C", "C", "C", "S", "A", "C", "C", "S",

"S", "S", "S", "S", "C", "C", "A", "C", "C", "C",

"C", "C", "S", "S", "C", "S", "S", "S", "S", "S",

"S", "C", "C", "S", "C", "C", "C", "A", "C", "C",

"C", "S", "S", "S", "C", "C", "S", "C", "C", "C",

"A", "C", "C", "C", "C", "S", "C", "C", "C", "S",

"C", "C", "S", "S", "C", "C", "C", "C", "S", "C")

liverwell_cat <-

c("S", "A", "S", "S", "S", "C", "S", "C", "A", "S",

"S", "C", "C", "C", "S", "A", "C", "C", "A", "A",

"A", "A", "C", "A", "S", "C", "C", "A", "C", "C",

"C", "A", "C", "C", "C", "C", "S", "S", "C", "S",

"S", "C", "C", "C", "C", "S", "A", "C", "S", "C",

"S", "A", "S", "C", "S", "C", "C", "C", "C", "C",

"S", "C", "C", "S", "S", "C", "C", "C", "C", "C",

"C", "A", "C", "C", "C", "C", "S", "C", "A", "C",

"A", "A", "C", "C", "A", "C", "C", "C", "C", "S",

"C", "A", "S", "S", "A", "C", "S", "C", "S", "C")

leporsy_foundation_cat <-

c("A", "A", "A", "A", "A", "C", "A", "C", "A", "C",

"S", "C", "A", "A", "S", "A", "C", "S", "C", "A",

"A", "C", "A", "C", "S", "S", "A", "C", "A", "S",

"S", "S", "A", "C", "C", "A", "C", "A", "A", "A",

"S", "A", "A", "C", "A", "S", "C", "C", "A", "C",

"A", "C", "A", "C", "A", "C", "A", "C", "C", "A",

"C", "A", "A", "S", "S", "C", "C", "A", "C", "A",

"S", "C", "C", "C", "C", "S", "A", "S", "A", "S",

"C", "C", "S", "A", "S", "S", "S", "S", "S", "S",

"S", "S", "S", "C", "A", "A", "S", "S", "C", "A")

movember_cat <-

c("A", "S", "A", "A", "A", "C", "S", "S", "S", "C",

"C", "A", "S", "S", "S", "C", "C", "S", "S", "S",

"S", "C", "C", "A", "A", "A", "A", "A", "S", "C",

"C", "C", "C", "C", "S", "S", "S", "S", "C", "S",

"S", "A", "S", "S", "A", "C", "S", "S", "S", "C",

"A", "S", "C", "A", "S", "S", "S", "S", "S", "C",

"S", "A", "C", "A", "S", "S", "C", "A", "C", "C",

"A", "C", "A", "S", "S", "S", "C", "A", "A", "C",

"A", "C", "S", "A", "A", "C", "A", "A", "C", "C",

"C", "A", "C", "C", "A", "C", "A", "C", "C", "A")

sj_nsw_cat <-

```
c("S", "C", "S", "C", "C", "S", "S", "A", "C", "C",  
"C", "C", "C", "C", "C", "C", "C", "C", "C", "C",  
"S", "C", "S", "S", "S", "S", "C", "S", "A", "C",  
"S", "A", "S", "C", "S", "C", "C", "C", "C", "C",  
"C", "C", "C", "C", "S", "C", "S", "S", "C", "S",  
"S", "S", "A", "S", "C", "S", "C", "C", "S", "C",  
"C", "C", "C", "C", "S", "S", "C", "S", "C", "C",  
"C", "C", "C", "S", "S", "C", "C", "C", "C",  
"S", "C", "C", "S", "S", "S", "C", "C", "C", "C",  
"S", "C", "C", "C", "S", "C", "C", "C", "C", "S")
```

```
vinnie_nsw_cat <-
```

```
c("S", "S", "C", "C", "S", "C", "A", "C", "C", "C",  
"C", "S", "C", "S", "C", "C", "C", "S", "S", "S",  
"C", "C", "C", "A", "C", "C", "C", "C", "C", "C",  
"C", "C", "C", "C", "S", "S", "S", "C", "C", "S",  
"C", "A", "C", "C", "C", "C", "C", "C", "A", "C",  
"C", "C", "C", "C", "C", "C", "S", "S", "A", "S",
```

```
"C", "C", "S", "C", "S", "C", "S", "C", "S", "C",
```

```
"C", "S", "A", "A", "S", "A", "C", "C", "S", "C",
```

```
"C", "S", "C", "C", "C", "C", "C", "S", "C", "C",
```

```
"C", "S", "S", "C", "C", "C", "C", "C", "C", "S")
```

```
deaf_nsw <-
```

```
c("C", "C", "C", "S", "S", "S", "S", "S", "S")
```

```
# add category columns to dataframes.
```

```
gd_nsw_cd[, "cat"] <- guide_dogs_cat
```

```
tl_dfnsnw_cd[, "cat"] <- deaf_nsw
```

```
tl_hf_cd$cat <- heart_foundation_cat
```

```
tl_lw_cd$cat <- liverwell_cat
```

```
tl_movember_cd$cat <- movember_cat
```

```
tl_sja_cd$cat <- sj_nsw_cat
```

```
tl_tlma_cd$cat <- leporsy_foundation_cat
```

```
tl_vnsw_cd$cat <- vinnie_nsw_cat
```

```
tl_ccnsw_cd$cat <- tl_ccnsw_cd$cancer_council_cat
```

```
massive_df <- bind_rows(gd_nsw_cd, tl_ccnsw_cd, tl_dfnsnw_cd, tl_dnsnw_cd, tl_hf_cd, tl_lw_cd,  
  tl_movember_cd, tl_sja_cd, tl_tlma_cd, tl_vnsw_cd)
```

```
# likes_photo <- massive_df[, c("media_type", "favorite_count")]
```

```
#g <- unlist(likes_photo[[1]])
```

```
#gg <- cbind.data.frame(g, massive_df$favorite_count)
```

```

# photo_df <- gg[gg$g == "photo", ]
# nphoto_df <- gg[gg$g == "no-photo", ]

#> var(c_df$favorite_count)
#[1] 31.46858
#> var(a_df$favorite_count)
#[1] 14.99052
#> var(s_df$favorite_count)
#[1] 63.73035
#> 54/15
#[1] 3.6
#> res.aov <- aov(favorite_count ~ cat, data = massive_df)
#> res.aov
#Call:
# aov(formula = favorite_count ~ cat, data = massive_df)

#Terms:
#          cat Residuals
#Sum of Squares   350.22 32906.16
#Deg. of Freedom    2    858

#Residual standard error: 6.192913
#Estimated effects may be unbalanced
#> summary(res.aov)
#      Df Sum Sq Mean Sq F value Pr(>F)
#cat      2   350  175.11   4.566 0.0107 *
#Residuals 858 32906   38.35
#---
#Signif. codes:  0 'b

```