

TUN Data Challenge

Hubspot and DonorPerfect

Sneha Jinturkar
Sayali Pendharkar
Kshitija Joshi

Abstract

The TUN data challenge is based on the program Hire Heroes USA, which connects transitioning veterans to their next careers. The aim of the challenge is to use the data concerned with the program and answer specific questions, based on analysis, visualization and examination of the data set(s). Our task was to analyze if there is any historical correlation between social media messaging or online presence and actual online donations. We worked on the Hubspot and DonorPerfect data sets, which had information from social networking sites - Facebook, Twitter and LinkedIn. Based on the campaigns, the online response and the funds raised, we generated reports on how these funds were raised, while individually analyzing the effect of each of the social networking sites.

Problem and Motivation

Problem statement:

To analyze historical correlation between social media messaging/online presence and actual donations and measurable effects on fundraising over the years/months from the brand awareness created by different channels of social media.

A proper understanding of the situation given to us:

Hire Heroes USA is a non-profit organization that empowers transitioning U.S. military members, veterans and their spouses to succeed in the civilian workforce through workshops, personalized career coaching and employment preparation counseling. All services are provided at no cost to the transitioning military member, veteran or spouse.

Our task is to find out if the social media presence affected the cause to move in a direction. To do this, we analyzed two data sets, one which had information on social media and the other on the funds received. We related the two using the one factor, time.

Approach

Below is the step-by-step procedure we followed to analyze the data:

- **Understanding the problem statement:** We tried to understand what exactly was the motive behind this analysis. We looked at what Hire Heroes was, what it did and how this data could have been relevant to it. We used two data sets to analyze this - DonorPerfect and Hubspot
- **Removing inconsistencies from the data sets:** We used Excel and R to clean the data. There were a lot of inconsistencies which could have misled our analysis.
- **Perform exploratory analysis using Tableau:** We used Tableau to relate and understand each of the columns in the data sets, both individually and together. The objective was to find out if any of the social networking sites had impacted upon the donations made to the campaign, both positively and negatively.
- **Report generation using Tableau:** Based on the analysis, we projected them on Tableau and generated reports.

Tools Utilized

The following were the tools we used in our analysis:

- **Excel:** Microsoft Excel is a spreadsheet developed by Microsoft. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications.
Tasks done using Excel -
Replaced missing values by NULL and filtered data by excluding NULL values.
Replaced 'Hire Heroes USA' as a generic account name for all channels.
Removed special characters appended for all dates in 'Donorperfect' data.
- **R:** R is a programming language and software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing.
Tasks done using R -
We used R in the pre-processing step to clean the data sets as per our requirements. The main issue we faced was with the date column. We had to relate the dates and make sure that we consider similar data which is available in both, DonorPerfect and Hubspot.
- **Tableau 9.2:** Tableau software produces a family of interactive data visualization products focused on business intelligence.
Tasks done using Tableau -
We used Tableau version 9.2 to analyze and represent the data in graphical format. We worked on making sure that we used colors that were not a problem to the people who are color-blind.

Results and conclusions

Comparison of how much campaigning over 2015-2016

Most of the campaigning done over the year 2015 had great response compared to response received over the year 2016. It is observed that 2015 was an eventful year with respect to the donations received. The maximum number of donations generated in the year 2015 were made by individuals.

The 'Holiday Campaign' and 'This Day In History' were most successful campaigns although those were not amongst most favourite posts on Twitter (correlation)

Most of the campaigns were published via Twitter. 'Holiday campaign' and 'This Day In History' were most successful campaigns through Twitter however these two were not the most favourite campaigns of Twitter

People responded and liked campaigning on social media however they have not donated much

'Registration' seems to be the most favourite campaign however it's not the most successful one(around 180k). 'Holiday Campaign' and 'GivingTuesday' received donations more than 200k.

'Throwback Thursday', 'Registration', and 'Training Programs' seemed to be always active in campaigning through social media

As it is observed that over 2015 lot of campaigning was done through various social media such as Facebook, LinkedIn, Twitter compared to campaigning done over 2016. It is observed that 'Throwback Thursday', 'Registration', and 'Training Programs' were one of those few campaigns published in 2016 and also well responded in 2015.

Overall time spent(percentages) and donations received for a particular campaign:

Analysed overall time spent over campaigning. Calculated '% Total time' by calculating difference between (Published time - Created time). 'GivingTuesday' and 'Holiday Campaign' were top two successful campaigns which received more than 200k. Substantial time spent on for these campaigns.

Analysis of campaigning conducted via Twitter for 'Registration' campaign:

LinkedIn likes and comment have made a significant effect on 'Registration' campaign. The donation received by this campaign was around 180k, for which LinkedIn must have made a significant impact. Also it can be seen that even if less time was spent(slide 16) on 'Registration' campaigning, it has received best responses over all social media

Conclusion: We observe that there is a high correlation between the social media presence and the donations received. The more the campaign was publicized, the more response it received in terms of donations.