2017/18
2016/17

# Data Driven Non Profit

## Importance, Challenges & Solutions

A NuSocia Publication

2017/18
2016/17

Graph / Statistic 3

Sequidel lacius exerspit, nes excescient hictior iatemquid ute liquam voles aliscit inulpa consectem aut a corem fuga. Edia volum volupit incillique il nihilicillo vent Hitem quamoccu

according to 2017 survey

NEXT

**AVIRAL APURVA**
Co-Founder & CTO, NuSocia

# Table of Contents

# Understanding Your Data

Non-profits and NGOs operate with limited resources and finances. They need to make sure that they are using their time and money as efficiently as possible. It is possible for Non-Profit Organizations to find it difficult to find ways to measure the success of their efforts. In this ebook, we will explore how data can be used by Non-Profits for various purposes such as fundraising, improving operational efficiency, or improving decision making.

Data can be used by non-profits in a variety of ways. Data can be used for fundraising, improving operational efficiency, or improving decision making.

Data can be used in many different ways. It can help donors understand the progress of their donations, it can be used to measure the effectiveness of a charity's work, and it can help nonprofits with their decision-making.

Donors want to know how their money is being spent and what the impact is. This is where data comes in; data helps them understand where their money is going and what they are funding. Data allows charities to measure the success of their work, like how much food they distribute or people they reach out to for medical attention. This information helps nonprofits make better decisions that will affect more lives for the better.

Data can be an incredibly valuable asset for non-profit organisations. It can help the organisations identify weak points in what they are doing, set better goals, and come up with new approaches to their work.

With access to data, nonprofits can see how many people they have reached or haven't reached. They also get to see which causes are more popular than others, where their money is being spent and who they should be targeting in the future. Data is key for any organisation to succeed in today's world because it helps them figure out what is working or not working.

# How to Incorporate Data Driven Process for Non Profits

When you're a non profit or NGO, it's hard to know where to start. You want to provide valuable services and do good work, but you also want to ensure that the funding is there so you can keep your doors open and your employees employed. Data driven processes can help you accomplish all of those goals: they will indicate how much money is needed, so you can increase fundraising efforts and be confident in what you need for operations, and they will help drive better service delivery since you'll have an accurate picture of who needs what services.

The first step in becoming a data-driven nonprofit is to see the value in your organization's data. Data isn't always factual or actionable, so be sure to collect it productively so you aren't left with an overload of information.

I would like to put forward a 5Cs process that if adopted can help non-profit power their initiatives with a data driven approach.

| Curate | Confirm | Collect | Compile | Communicate |
|--------|---------|---------|---------|-------------|
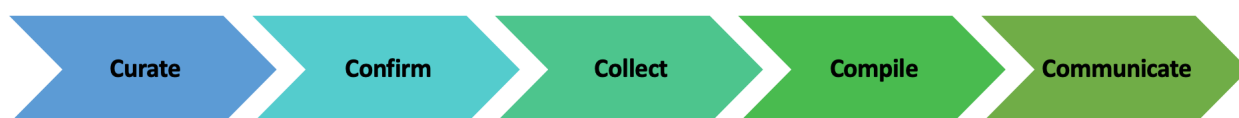
- **Curate**- This stage is a starting stage where the team has to ask and curate the right set of questions which are easy to understand and captures the maximum information on ground. This is the stage where the team should spend substantial time and resources to make sure that the questions that are being curated are of high-quality and something that captures the context of the initiative on ground and eradicating biases as much as possible.

- **Confirm**- Once the right questions have been curated by the team they should not hurry and start collecting data based on that but rather spend some time confirming the integrity and the output of the questions that are curated by doing a trial and a mock run with the questionnaire and checking for any technical challenges beforehand.

- **Collect**- This is a stage where actual data collection would be done on ground with the set of right questions which both the data collected as well as the respondent understands and can answer with clarity. Various standardized approaches should be set up at the stage of data collection to make sure that the data that is getting collected is having a standard format throughout.

- **Compile**- the data that is getting collected on ground would be coming from various sources so it's very important to take in consideration of the step of setting up a process to compile this data such that there is minimum or no loss of information as well as all the sources are playing nice with each other and resulting in a standardised format of the dataset. This will help create the process of Incyte generation and analysis of data easy and more accurate.

- **Communicate**- Data collection and analysing is nothing without effective communication and storytelling around it. To become a true data driven organisation nonprofits should look into generating a capacity around storytelling in a way that they can reflect the real impact that is being generated on ground which can be backed by the facts and data that has been generated and collected.

# Issues and Challenges With Data Collection on Ground.

There are many issues that NonProfits and NGOs face when collecting data. How does one know that the data they are collecting is accurate? What types of data should they be collecting? How do they securely store their collected data? These are all questions that need to be answered for NonProfits and NGOs in order for them to be competitive within today's society.

Let's take a look at some of the challenges that the non-profit face on the ground:

- **Inconsistent data collection**: generally the organization does not put enough effort into understanding the various standards that are needed to be kept in mind while collecting the data on ground. This is very important because if the data is not coming out consistently it will definitely cause an issue when the actual analyzing of the data will be done. Inconsistent data can lead to added efforts at the time of data cleaning and making actionable insights from it. This is something which the majority of the time is overlooked but it should be something that is put in process from the day one of data collection.
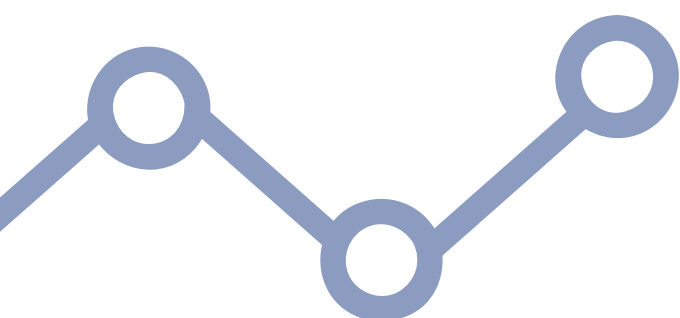
- **Context of data collection**: this type of challenge happens all the time specially in the social sector because when the project would have been defined it would have been created with a different context and design in mind. But when the Data collector goes on the field and tries to collect the data and his or her perspective comes into play and it's very important to take care of this bias because if the context of the data getting collected is not aligned with the actual implementation of the project then the outcome of this kind of data would not be that much helpful and insightful towards actions.

- **Complexity**: one thing to note here is that the data collection has to be as simple as possible. At the time of designing the data collection tool as well as while capturing the team should try and maintain the complexity of the questions to a very basic level so as to capture more insights when the respondent is actually able to understand the question because it's much simpler and the complexity is also lower.

- **Lack of training in data collection**: now this is something which the majority of the nonprofits face because of the lack of resources and sometimes negligence. Training of data collectors is very important such that the person who is going on ground to collect the data understands the challenges and the context of the problem for which they are trying to collect the data. Training also involves all the possible technical and IT related challenges that the collector can anticipate when they are going on the field and have a solution ready when those issues arise. Lack of training in data collection can easily result in low quality of data collection because proper processes were not being followed and if there are multiple collectors working together there would be an issue of consistency between their work due to lack of training.

- **Economic and IT restrictions**: this is also very important because sometimes there are certain logistic restrictions and challenges that one need to always anticipate before going into the field. For example, the majority of the villages still don't have internet connection and even electricity to charge simple devices. The team which is going on the field to collect the data has to account for all of these challenges and be prepared for it so all the IT related restrictions and logistical challenges that can arise should be taken care first hand before going in the field.

- **Lack of quality assurance processes**: finally setting a quality assurance process is very important when the data collection is happening because in jollity of the projects there are multiple stakeholders as well as data collectors that are involved and keeping a consistency among them so that data and that is getting collected is of high integrity certain quality assurance processes have to be incorporated. This can be as simple as finalising a single tool that the whole team will use while collecting the data rather than different people using different tools or methods to collect the data.

# Importance of Data Quality Control

Data quality control is an important process in any company or organization because it ensures that the data being used to run the business are accurate and up-to-date. Poor data quality can cause errors in operational processes, which can have negative consequences for organizations.

Data quality control is integral for non-profit organizations and NGOs. Data comes in many forms - statistics, reports, statistical models, data sets, among others. And it's one of the most under-appreciated aspects of digital transformation. The best way to get accurate data is to put in the time and effort to ask the right questions before you start collecting them.

There are many steps that need to be taken in order to ensure data integrity. One way is to create guidelines on what can be done with data, like not using it for anything other than for how it was meant. You need to continually verify your data and always work on improving it.

One of the problems associated with data quality control is that the reliability of your data may be called into question. In a recent study by Nonprofit Hub, 90% of the non-profits said they collect data, but a surprising 49% said they don't know how data is collected. Not knowing how data is collected can lead to problems with proper analysis and the quality of the stored data.

Reliability is a key term when data is used in decision-making. If data is unreliable, you run the risk of making bad decisions that could in the long run hurt the organization.

Compatibility is a major pain point for nonprofit professionals when it comes to the management of data. The most glaring problem with data collection is that nonprofit organizations are likely to work with multiple software solutions that do not play well with each other. For example, with multiple tools and systems, certain data can be collected from specific donors, but they are not designed to communicate with each other because the data has to be put together into a single format which becomes difficult due to it being coming from multiple different sources.

Creating and maintaining a data quality system is a necessary task for any organization. A data quality audit will show you what needs to be done to ensure that the data is accurate and relevant. Checklists, metadata, and clear labeling of the data are just some of the ways that these issues can be addressed.

All in all, when properly leveraged, data can be the ultimate tipping point. When properly collected and stored, it can help to review campaign performance, provide a complete overview of the life cycle of your donors, improve communication, and make important decisions for the entire organization. Many nonprofits collect data, but most don't get it right.

# Tools for Data Collection

As we know that nonprofits operate on limited resources both from manpower as well as financial constraints, it is very important to keep in mind to opt for tools which are taking into consideration these contingencies and are providing the maximum efficiency in the process.

There are many tools which are available for free of cost like a software as a service product from various companies and some of the projects which are open source can also be utilised by nonprofits as it is free to use for all. Let's see some of these tools below:

- **Google Forms**: Google form is a simple yet very powerful tool for data collection if used wisely. It is an offering from Google in the form of a form that one can create very easily without having any technical knowledge. It is also available for free to use by anyone and only a gmail account is required to access it. When you use Google form, it will collect and organize all of the data in a spreadsheet which can be accessed and added to easily. You can also share the survey with colleagues, partners or random people that you want to get their opinion on the project. One thing to note with Google form is that to input the data it requires you to be connected to the Internet so in some of the cases it might not be the right tool to pick for data collection in areas where there is an issue with network connectivity. But due to its robustness freely available it is surely one thing that nonprofits can look at if the above mentioned drawback is not present.

- **Excel/Spreadsheet:** Another tool that can be used by the researchers on ground is Excel. Although it would not be the most elegant solution but if there is a constraint of finances then this is something that can be taken into consideration. The good thing about XL is that it's now available on mobile devices in the form of an application and can be used freely to input the data. A use case could be when the team is going on field people can start collecting the data in their own application having the same columns designed in the form of questionnaire and once the data collection has been done by all the collectors a single Point of contact can merge all of these into a single spreadsheet or Excel to get the final dataset. This has one additional step to merge the data from various sources manually but it takes care of collecting the data off-line without any network issue and for free of cost.

- **KoBo toolbox:** this is also one of the tools which can be used to collect data even in scenarios where low connectivity is present. This tool is available for free for teams to create and publish forms and capture the data on apps that are available for it and capture data on ground and the data will be getting collected at a single source which can be then easily exported in the form of excel file or a CSV file. This tool has a bit of a learning curve to it but it does not require you to be technically sound if a little bit time is spent on the platform to understand various options even a non-technical person can create and publish forms very easily. This is a good alternative to per data collection platforms as it has its own app and the data is getting stored in a cloud based storage which can be accessed from anywhere and can be downloaded at any time.

- **Open Data Kit:** Open data kit is a tool which is an open source and freely available tool for data collection in areas where low or no network connectivity is available and a single storage centre for all the data collectors collecting data on ground. This tool being open source means that the source code of this tool is freely available for everyone to use and modify according to their liking. The only contingency with most of the open source tools is that the person using these tools is the only person liable for making it up and running on their own servers and then maintaining it if something goes down. So this can be a good option for organisations which have technical capacity and can deploy this tool to their own servers and maintain it.

# Data Cleaning - What, Why & How

Data cleaning is a very crucial and important step that comes after the data has been collected and before actually going and digging deep into the data to analyse it and generate insights from it. Data cleaning is important because in the real world the data that gets collected is filled with noise. Noise means that there are a lot of issues with the data for example missing values where some of the data points are missing. Redundant values where some of the data has been input more than one time creating a redundancy of the data point within the dataset. Another one could be an incorrect value that has been input in a data point which requires a specific data type.

In real life these are the challenges that come with the data whether we want it or not. There are other issues that can also create challenges to get accurate insights, could be the outliers that are there in the data which generally are not following the trend of the data set. Data cleaning is a very crucial step because it takes care of all the above mentioned issues with the data such that the data is free from all of these noise and issues and the insights that now will be generated on this data set is of high integrity and accuracy.

There are various tools and techniques that are available for data cleaning some of them are listed below:

- **Excel** : Excel is a very powerful tool which should not be undermined and should be leveraged properly to utilise all its features. Excel has a lot of features which helps us to take care of cleaning the data. For example if there are some data points which are missing excel can help us find these data points all at once and replace it with a universal value to streamline the data. This is just one example but Excel has a lot of options which if learnt can help in majority of the cases in data cleaning.

- **Open Refine**: Open Refine is an open source tool initiated by Google. This tool is a GUI-based tool which can be used to clean out the data having the above said issues within it. Using Open Refine is very easy and user-friendly because of its Dragon dropped Best functionality where the users can just import the data set and can just click on the options of various data cleaning techniques that they need to run on their data set. This is a very intuitive free to use tool that the nonprofits should definitely take a look at.

- **Python/R Programming:** These two programming languages are also something which can be used to handle and do the data cleaning because of numerous libraries that are available for these programming languages for this task. Using a programming approach to clean the data can be very helpful in the case where there are some advanced and customized approaches that need to be applied on the data to clean it.

This is because of the nature of programming which gives the full flexibility to the user to code whatever they want to do with the help of processing power of the computer. The only drawback of programming languages and its usage and data cleaning is the technical expertise that is required to do these tasks. It is recommended for the non-profits which have technical bandwidth within their organisation and operate on a large dataset which requires customised and advanced data cleaning techniques.
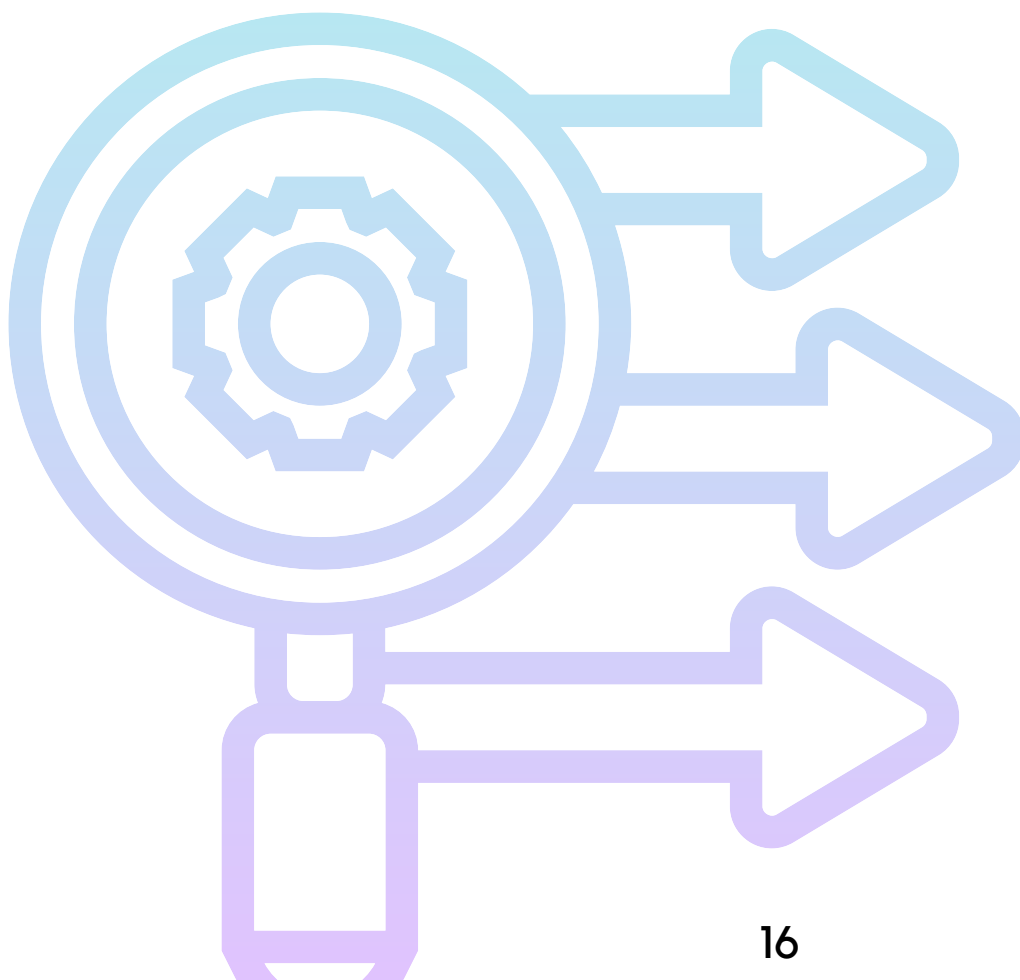
# Road Ahead

Artificial intelligence is something that is a hot word around the globe in almost all the industries right now. And also in the social sector there are instances where AI for social good has been utilised to create impact that is evident and on a large scale. AI nowadays is being used to tackle specific challenges on ground for example helping farmers eradicate pesticides from their crops utilising AI techniques by predicting it. Drones are now being used to capture the Geo-spatial data of certain initiatives on the field. AI is also being used to gather the data on field using natural language processing and then later converting it in the form of a data set. So, It is very important for nonprofits to at least start looking at what the technology has to offer and how it can be incorporated with the innovative approach to bring efficiency in their workflow on field and in their operations. Needless to say all of these things have the same backbone which is data. Even techniques like machine learning, data science and AI all rely on data.

Thus it is very important for nonprofits to understand the value of the data that they are generating and capturing on the field and make sure that the data that has been collected by them is of utmost quality and have high integrity such that if they want to apply these advanced techniques to their data to bring some innovative and actionable insights for decision making it is coming out with at most accuracy and high impact insights.