CLARITY
HUMAN SERVICES

HMIS Data Visualization

Using Data Visualization to Achieve Measurable Results

White Paper

July, 2015

Bitfocus

1. EXECUTIVE SUMMARY

Understanding homelessness – its antecedents, causes, and characteristics – is critical to its elimination.

The only way to fully understand homelessness is to glean data from the Homeless Management Information System (HMIS).

HMIS data is indeed solving the problem of homelessness in America. According to the *2015 The State of Homelessness in America*, released by the National Alliance to End Homelessness:*

✤ 34 U.S. states had decreases in overall homelessness between 2013 and 2014.

✤ 40 states had decreases in the number of people living in unsheltered locations.

✤ The majority of states had decreases in every major subpopulation: family homelessness (32 states), chronically homeless individuals (27 states), and veteran homelessness (28 states).

✤ Rapid re-housing capacity grew dramatically—nearly doubling from 19,847 beds in 2013 to 37,783 beds in 2014, a 90 percent increase. 40 states increased rapid re-housing inventory.

✤ The number of permanent supportive housing beds continued to grow from 2013 to 2014 by 15,984 beds (5.6 percent) to a total of 300,282 beds. 35 states saw increases and 15 states saw decreases.

Our work is not over yet, however. For example, the same report also noted that 17 states saw increases in

Stories of Homelessness

Each piece of HMIS data is telling a story – it's explaining how and why certain service structures are more effective than others for certain populations. It's telling the story of the routes that people followed that led to their homeless, unveiling their vulnerabilities and unfortunate circumstances along the way.

The Importance of the HMIS

Understanding these stories is the only way we can create the programs and services necessary to end homelessness. The rich data sets stored in the HMIS can help you to learn about the clients you serve, and thereby develop more effective policies and programs to meet their needs.

---

* As of the publication of this paper, this is the most recent official data available.

homelessness between 2013 and 2014. Unaccompanied youth and children accounted for an unacceptable 7.8% (45,205 youth) of the total homeless population, and nearly 50% of these youth were residing in unsheltered locations. These numbers emphasize the need to use the HMIS to its fullest capacity.

Measurable Results

Generating information by transforming variables into measurable factors. This process defines 'fuzzy' concepts and allows them to be measured empirically and quantitatively. This information is then applied to enable informed decision-making.

The HMIS is the critical conduit that funnels and processes the tireless efforts of homeless assistance providers, churning out the reports that accurately inform policy-makers at the local, state, and Federal level.

In short, HMIS data is the only way to develop policies, procedures, and programs based on measurable results.

Within the HMIS lie countless stories of homelessness. Each client has a story, and the HMIS codes these stories into information that we can act on.

Understanding these stories is the only way we can create the programs and services necessary to end homelessness. The rich data sets stored in the HMIS can help you to learn about the clients you serve, and thereby develop more effective policies and programs to meet their needs.

However, HMIS data, regardless of its level of quality, is not useful unless it is understood. As mentioned, each piece of HMIS data is telling a story – it's explaining how and why certain service structures are more effective than others for certain populations. It's telling the story of the routes that people followed that led to their homeless, unveiling their vulnerabilities and unfortunate circumstances along the way.

The value of HMIS data is null unless it is presented in a digestible, understandable format.

Therefore, the secret to understanding these stories, and transforming their lessons to into measurable results does not lie solely with HMIS data. Instead, measurable results are fully dependent upon how the data is presented – or in other words, it is dependent upon data

visualization techniques used to present the data. *Data visualization is key to accurate and understandable storytelling.*

Oftentimes, stories will be hidden within HMIS data, undetectable by even the most sophisticated reports. Many communities and CoCs have 'hunches' regarding the patterns they believe lie within their data, but these patterns remain stubbornly hidden within the data. For example, a community may have a hunch, based on day-to-day client interaction, that a program's recidivism rate is higher for adult clients struggling with substance abuse, but substantially lower for homeless youth. However, their standard reporting measures do not substantiate this belief.

Data visualization is their secret to uncovering this particular hidden data trend, as well as other hidden 'gems' that lie untapped within their HMIS.

Purpose

The purpose of this white paper is the following:

- Define Data Visualization and its importance to homeless assistance
- Identify the characteristics of effective data visualization
- Understand the basics of data visualization techniques
- Present examples of data visualization and HMIS standard HUD reporting

Table of Contents

2. DATA VISUALIZATION
A Picture is Worth a Thousand Words

2.1 Data Visualization: What It Is and Why It's Important

2.1a) What is Data Visualization?

Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context.

Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization techniques.

Data Visualization

Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context.

In short, the primary goal of data visualization is to clearly and efficiently present information using statistical graphics, plots, information graphics, tables, and charts that have been strategically chosen to portray the particular data at hand.

Two hallmark benefits of data visualization are *exploration* and *explanation*:

Exploration:

Find and understand the story that the data is telling you. The concept of **understanding** is particularly important here – data visualization gives you the tools you need to not only detect the data's story, but also understand this story's meaning and relevance.

Explanation:

Tell this story to the audience in a way that is easily digestible and understandable.

Other benefits of data visualization include the ability to represent large quantities of data coherently, and discern relationships within this data without distorting the data itself. It makes complex data more accessible, understandable, and usable, allowing the audience to easily understand patterns of importance.

2.1b. Characteristics of Effective Data Visualization

In his book *The Visual Display of Quantitative Information*, Edward Tufte states that data visualization should:

- ✤ Induce the viewer to think about the story of the data rather than about methodology, graphic design, the technology of graphic production or something else
- ✤ Avoid distorting what the data have to say
- ✤ Present many numbers in a small space
- ✤ Make large data sets coherent
- ✤ Encourage the eye to compare different pieces of data
- ✤ Reveal the data at several levels of detail, from a broad overview to the fine structure
- ✤ Serve a reasonably clear purpose: description, exploration, tabulation or decoration
- ✤ Be closely integrated with the statistical and verbal descriptions of a data set.

Data visualization must be approached with care, as inappropriate display choices can distort reality.

Several common data visualization mistakes include:
- ✤ **Inconsistent or reversed scales**
- ✤ **Too much information**
- ✤ **Encoding quantitative data incorrectly**
- ✤ **Proportional axis scaling**
- ✤ **Using counts vs. percentages when comparing periods with different totals**

3. DATA VISUALIZATION TECHNIQUES
Finding the Best Visualization Strategies for Your Data

3.1 The Basics – The 3 'What's' of Data Visualization

1. What Are the Characteristics of Your Data?
What is the size and cardinality of your data? Cardinality is the uniqueness of the data values contained in a data set. High cardinality indicates a large amount of unique values. Low cardinality means a set of data contains a large percentage of repeat values (i.e. gender or yes/no veteran status). Here are several examples:

Client Record Unique Identification Numbers – This would be high cardinality as each item is unique.

Client Gender or Veteran Status – This would be low cardinality as there will be a high amount of repeat values.

2. What Do You Want To See?
Determine what you are trying to visualize with your data. Are you trying to view the demographics characteristics of your data? Are you trying to understand causation or causality? Or perhaps you are trying to determine change over time. Understanding what you want to know will help you to know the steps to take to get there (e.g. what chart(s) to use).

3. What Are The Characteristics of Your Audience?
Who will be digesting your data? Will it be CoC board members? A collection of stakeholders? Or perhaps you will be presenting data at a community meeting?

It's also important to ask: What information does the reader need to be successful? And similarly, how much detail does the reader need?

Knowing your audience, and understanding how they typically process visual information, will enable you to choose the visual representations of your data that best convey the information in the simplest form for your audience.

3.2 Your Tools: What Chart to Use and When To Use It

Once you answer these questions, you're ready to pick the tool that will best visualize your data. Tables are great for sorting and looking up information, but graphs allow you to see patterns in your data, which is critical to understanding your data.

Other Data Visualization Design Considerations

....

✦ Avoid meaningless variety
✦ Avoid overwhelming gridlines
✦ Use color strategically
✦ Use text sparingly
✦ Avoid too much information

Time Series:

Do you want to show how your data has changed over time?

Column Chart

Use a column chart to show data changes over time **in general.**
In addition to regular column charts, there are two additional column charts that can help you fine tune your data visualization.
  ✦ Clustered – use when your data is not in any specific order (e.g. geographic names)
  ✦ Stacked – use to show the relationship of individual items to the whole.

Line Chart

Use a line chart when you want to emphasize ***trends***, or changes in data over time, rather than simply the amount of change (column chart). This chart is particularly helpful when understanding trends, or changes in data, is valuable to explaining your data.

Stacked Line Charts

Stacked line charts allow you to show the trend for multiple series of data. Note that these charts are not effective for showing comparisons between data sets. However, if the values on the X axis are continuous, a simple line chart is the best choice.

Area Chart

Use an are chart if you what to highlight how ***each*** of the values change over time.

Point In Time:

Do you want to show the current state of your data?

Bar Chart

Use a bar chart to show the values of several items at a single point in time. Bar charts are most commonly used for comparing the quantities of different categories or groups. Each value is assigned its own bar, and these bars can be configured either horizontally or vertically. The length or height of each bar represents the value.

As mentioned, bar charts can be vertical or horizontal. The horizontal format of bar charts make it easier to compare single point-in-time values than using a column chart. Also, horizontal bar charts allow more room if you have a long series of labels.

TIP: To help provide even further visual variance, assign each bar a different color.

Pie Chart

Pie charts show composition (how categories represent parts of the whole), and allow you to chart one data series in a small, succinct manner. If you want to highlight one particular value you can use an exploded pie chart.

TIP: It can be difficult to compare sections of a pie chart that are similar in size, but located in different areas of the pie, as opposed to next to each other. Therefore, it is most effective to include text and percentage labels within the chart itself.

Relationships Among Data:

Do you need to show relationships (i.e. similarities, differences, causality, correlation, etc.) among pieces of your data set?

Scatter Chart

Scatter charts allow you to show similarities (as opposed to differences) between large sets of data, and also allow you to show data grouping. These are most effective when there are many data points on the horizontal axis, or the values are not evenly spaced. In a scatter plot, each marker (symbols such as dots, squares and plus signs) represents an observation. Once you have plotted all of the data points using a scatter plot, you are able to visually determine whether data points are related. Scatter charts can also show correlation.

For example, mental illness is often correlated with length of homelessness; a scatter chart is a good way to visualize these types of relationships within your data.

TIP: Scatter plots are most useful for many data points. If you are working with only a small set of data points, a bar chart may be a more effective way to view the patterns in your data.

Bubble Charts

A bubble chart is similar to a scatter chart, but each bubble replaces the marker in a scatter chart. The location of the bubble represents the value for two measured axes (similar to a scatter chart), but the size of the bubble allows you to give additional context to the data value. For example, a large bubble can mean more data for that particular value in comparison to a smaller bubble for a different value in the same chart. Bubble charts are particularly helpful to visualize hundreds of values, especially if they differ in magnitude.

TIP: Animated bubble plots are a good way to display changing data over time.

Heat Map

Use a heat map when you need to use color to visualize large amounts of multi-dimensional data. Clusters of rows with similar values are displayed as areas of similar color. The easiest way to understand a heat map is to think of a table or spreadsheet which contains colors instead of numbers. The default color gradient sets the lowest value in the heat map to dark blue, the highest value to a bright red, and mid-range values to light gray, with a corresponding transition (or gradient) between these extremes.

3.3 About the Data Visualization Dashboard

A data visualization dashboard is a visual display of the most important information that you or your audience needs to know for a specific purpose. This information is consolidated and arranged on a single screen so the information can be reviewed at a glance.

Dashboard Design Tips:

- ❖ Keep all data presentation within the boundaries of a single screen
- ❖ Supply adequate context for the data
- ❖ Express measures directly
- ❖ Choosing appropriate media of display
- ❖ Apply only the necessary amount of variety
- ❖ Use high quality display media
- ❖ Encode your quantitative data accurately
- ❖ Highlight only what is important
- ❖ Do not clutter the screen with unnecessary decoration
- ❖ Use color strategically – and do not use it in excessive amounts

Dashboards are meant to be monitored at a glance, so it is important to keep context items near each other.

4. WHAT'S THE STORY?
Using Data Visualization to Get Measurable Results

HMIS data is the only route to obtaining measurable results. In turn, these measurable results are the only way to end homelessness in your community.

But how can you bridge the gap between raw HMIS data and measurable results?

The key is to know what types of stories are residing in your data. Once you understand these stories, you can select those that will portray your successes, failures, and needs in ways that are easily digestible to your audience.

Below are the most important clusters of stories that are lying dormant in your data right now. Using data visualization to explore and explain this data will pinpoint your community's needs so that a solution can be generated.

Stories of Client Characteristics and Service Needs

- ❖ How many people experience homelessness over the course of a year?
- ❖ What are the characteristics of homeless people, or those at risk for homelessness?
- ❖ Who in the community is most at risk of becoming homeless?

Stories of Service Utilization

- ❖ What are the average lengths of stay, and what are the contributing factors?
- ❖ What is the average number of homeless episodes?
- ❖ Who is utilizing each type of shelter?
- ❖ What is the prevalence of chronic homelessness? Veteran homelessness? Youth homelessness?
- ❖ What are the patterns of service use over time?

Stories of Program Effectiveness and Outcomes

- ❖ Which programs are effective and why?
- ❖ What are readmission rates for each program?

HMIS data can answer all of these questions and more. However, it can do so only through effective data visualization.

Answering these questions, and applying the proper data visualization techniques will enable you to tell the story you need to tell in a way that is interesting, and more importantly, understandable to your audience. It will also help you to detect those hidden stories that lie within the HMIS, stories that oftentimes have a substantial impact on the amount of funding a community receives, or does not receive.

When beginning a data visualization task, remember to consider these questions:

1. *What type of data am I working with?*
2. *What do I want to show with this data?*
3. *Who is my audience?*

5. CONCLUSION
Data Visualization to End Homelessness

Although CoCs differ in their structural organization and HMIS software applications, they share one nationwide commonality: they can use the data stored in their HMIS to improve their homeless assistance programs and better serve homeless persons and those at risk of becoming homeless. This means that every CoC has the opportunity to explore their data using data visualization techniques, which, in turn means that the benefits of data visualization are available to all CoCs.

Arguably, the most beneficial result of data visualization is the ability to detect hidden data trends that reside untapped within your HMIS. These hidden data trends can shed valuable light on the status of homelessness in your community, and they are only accessible through data visualization.

The most exciting aspect about using data to obtain measurable results is the fact that technology is swiftly advancing. This fact provides hope and excitement regarding the role of HMIS data visualization in the fight against homelessness.