

# Social Service Agency

Andy Grogan-Kaylor

January 02, 2021

## Contents

0.1 Background . . . . .	1
0.2 Clients . . . . .	1
0.3 Donors . . . . .	3

## 0.1 Background

You are working for a social service agency that keeps track of its clients and donors using Excel. While the agency has kept track of these two groups of individuals, the agency has done little to evaluate their work, and to assess ways in which they might do better.

The agency would like you to provide them with some data visualization, and they will provide you with a file of **clients** (in 2 formats)<sup>1</sup>, and a file of **donors**.

## 0.2 Clients

Here is what the data file of clients looks like.

NB that the mental health instrument is normed such that **100** is considered to be an *average* or *usual* mental health score. Higher scores indicate higher psychological wellbeing.

---

<sup>1</sup>This is an **advanced** topic, but it is worth thinking the idea that sometimes data visualization problems are really problems of how the underlying data is structured. Initially, the data called **clients** is easier to use, but for some purposes **clients\_longform** is easier to work with. The two files contain the same *information*. Are there circumstances in which one would be easier to work with than the other?

Table 1: Table continues below

ID	age	gender	program	mental_health_T1	mental_health_T2
1079	35	Female	Program B	87.09	102.3
2639	31	Male	Program C	85.47	86.26
2748	21	Male	Program B	98.22	109.2
2993	30	Male	Program C	87.66	94.18
1778	40	Male	Program C	88.29	104.8
3522	19	Male	Program C	80.49	92.57

latitude	longitude
42.33	-83.75
42.15	-83.75
42.23	-83.75
42.12	-83.81
42.19	-83.78
42.35	-83.66

The agency also had an intern from a school of social work who did some preliminary data visualization.

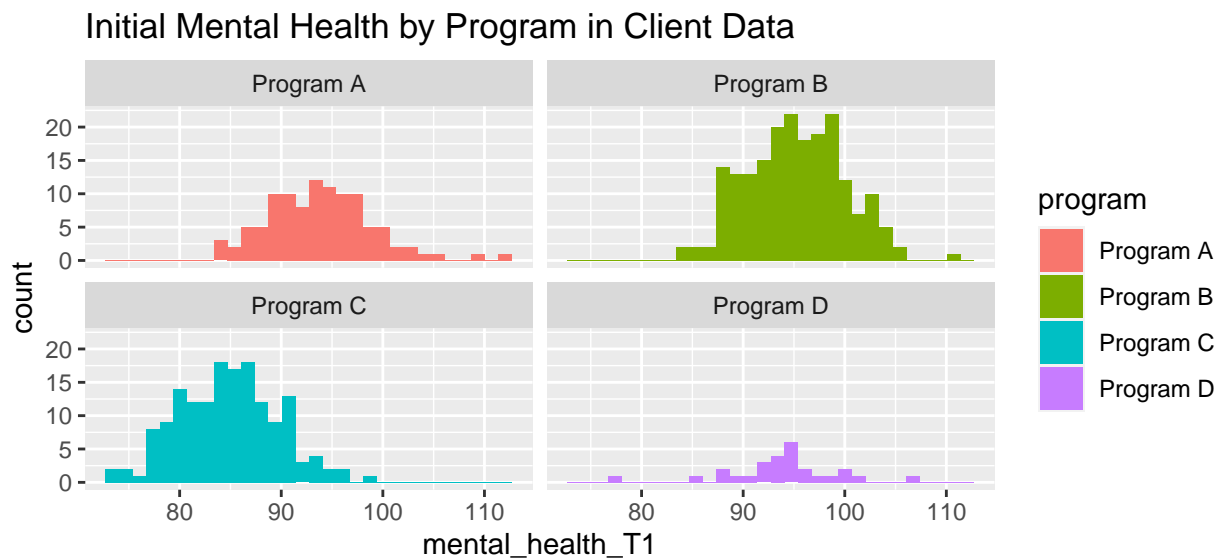


Figure 1: Preliminary Data Visualization of Clients

Agency staff would like to investigate questions of the following nature.

- What are demographic characteristics of clients?

- Do any demographic characteristics of clients have an association with mental health or program use?
- Is there any additional information that the agency should be collecting about clients?
- Are any of the programs associated with better mental health at Time 2?

The Director of the agency says:

“I’ve heard you can investigate things like this with `ggplot()`, maybe using something like `geom_point()`, `geom_smooth()`, and `facet_wrap(~)`.”<sup>2</sup>

### 0.3 Donors

The agency also has a file of donors which looks something like this:

ID	age	gender	neighborhood	annual_donation
NPcbuZh	44.06	Male	Neighborhood A	126
OX0x6Fv	39.34	Male	Neighborhood A	56
MdQQ1t8	48.05	Other Identity	Neighborhood A	230
FiNotKH	55.1	Female	Neighborhood B	1073
6Lp4W86	33.36	Other Identity	Neighborhood A	379
YLArlde	40.69	Female	Neighborhood A	289

The agency would like to know.

- In terms of available information, what are the demographic characteristics of donors?
- Is there any additional information that we should collect about donors?
- Is there any information that predicts which donors will donate larger amounts of money?

The agency did hire a consulting agency to do some preliminary analysis of their donors. After paying the consultant several thousand dollars they received a short report and the attached graphic.

The Director says:

“While the attached graphic may give you, and us, some guidance, we are wondering if you can produce anything that is a little simpler, a little more straightforward, and a little more clear.”

<sup>2</sup>Yes, the Director really talks like this sometimes.

## Donor Analysis



Figure 2: Preliminary Visualization of Donors