Social Service Agency

Andy Grogan-Kaylor

April 23, 2020

Contents

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), and a file of **donors**.

0.2 Clients

Here is what the data file of clients looks like.

Table 1: Table continues below

ID	age	gender	program	mental_health_T1	mental_health_T2
1089	20	Male	Program A	101.3	94.62
2012	20	Male	Program A	98.73	99.99
4384	31	Female	Program C	78.86	89.99
1764	25	Female	Program B	98.24	103.1
2560	19	Female	Program B	94.81	105.7
1497	30	Male	Program C	91.12	91.56

latitude	longitude
42.24	-83.57
42.2	-83.78
42.24	-83.85
42.33	-83.72
42.32	-83.75
42.31	-83.93

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

```
library(ggplot2)

ggplot(clients,
       aes(x = age, # x is age
           fill = gender)) + # fill is gender
  geom_histogram() + # use histograms
  facet_wrap(~gender) + # "small multiple" graphs
  ggtitle("Age by Gender in Client Data")
```
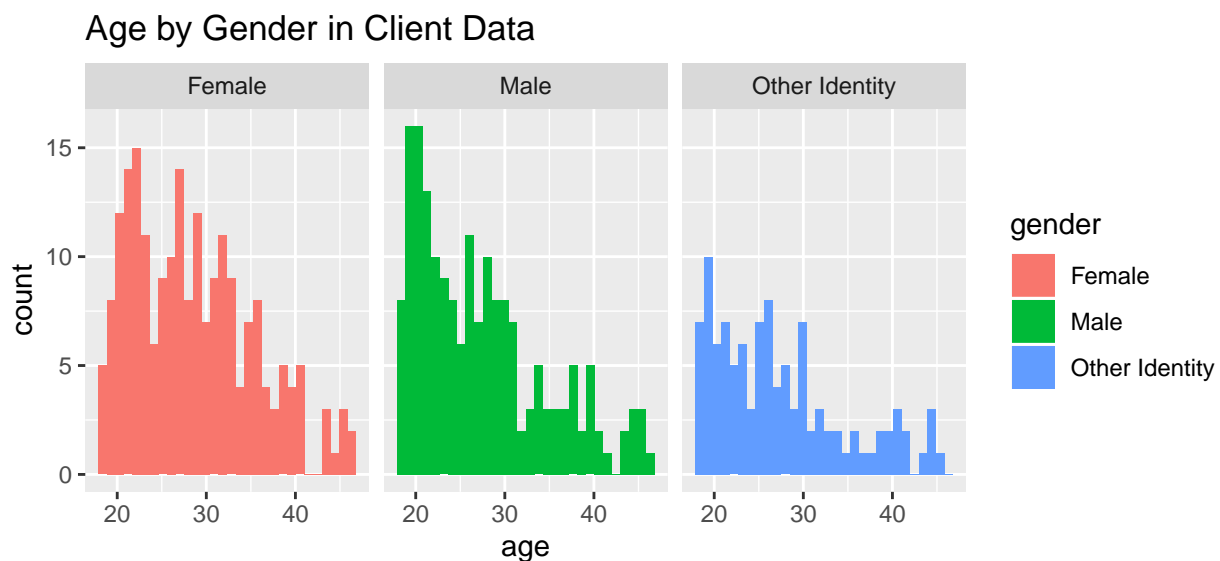

Figure 1: Preliminary Data Visualization of Clients

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(~)`."[1]

0.3 Donors

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

ID	age	gender	neighborhood	annual_donation
0h0Mnsh	41.58	Male	Neighborhood A	419
cyfyghu	47.05	Male	Neighborhood A	386
UZzIKhE	30.03	Female	Neighborhood A	85
EAslcJ1	36.2	Other Identity	Neighborhood A	216
D3jsJiP	35.53	Female	Neighborhood B	320
8GmBNo2	33.09	Female	Neighborhood A	339

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."
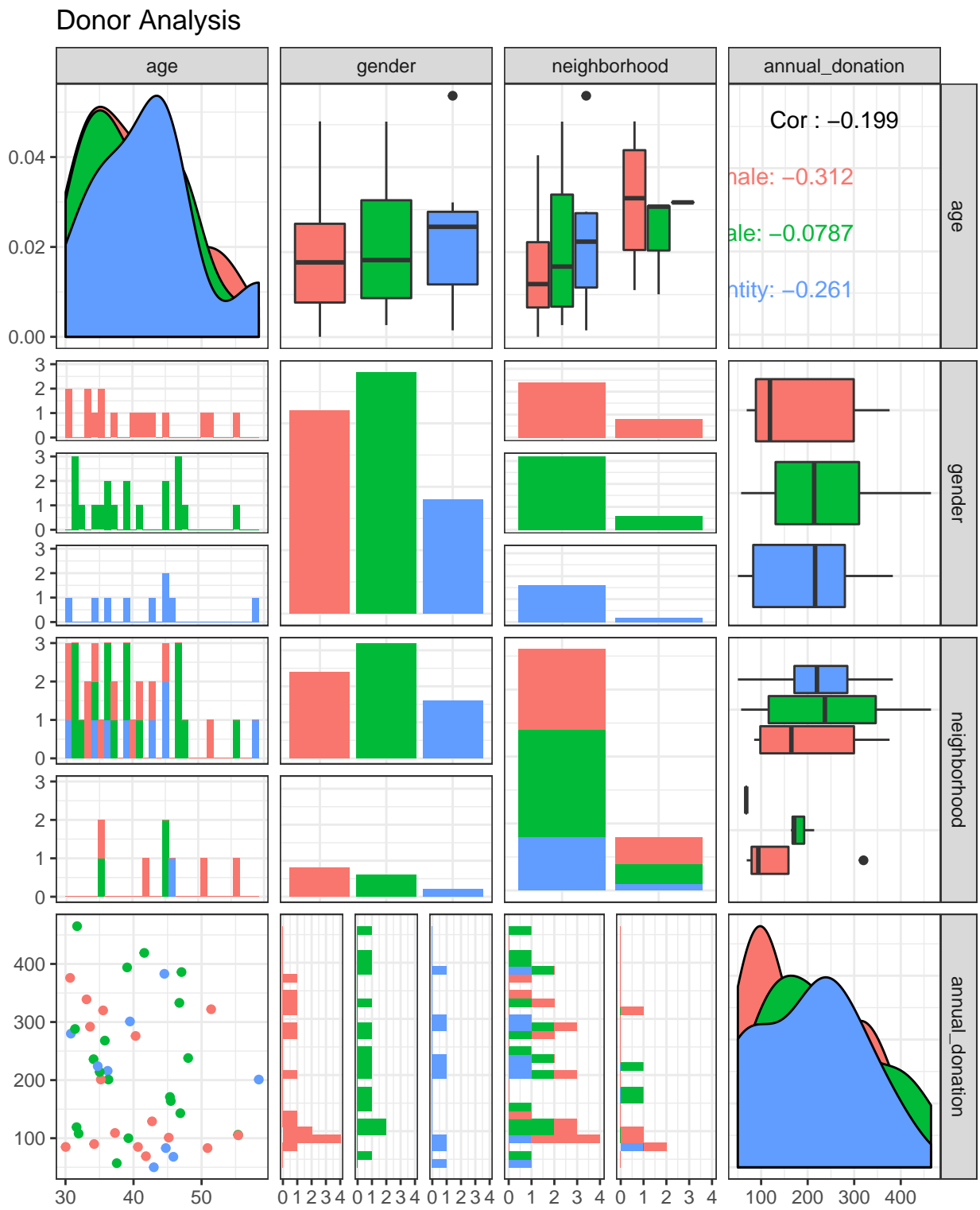
---

[1]Yes, the Director really talks like this sometimes.

Figure 2: Preliminary Visualization of Donors