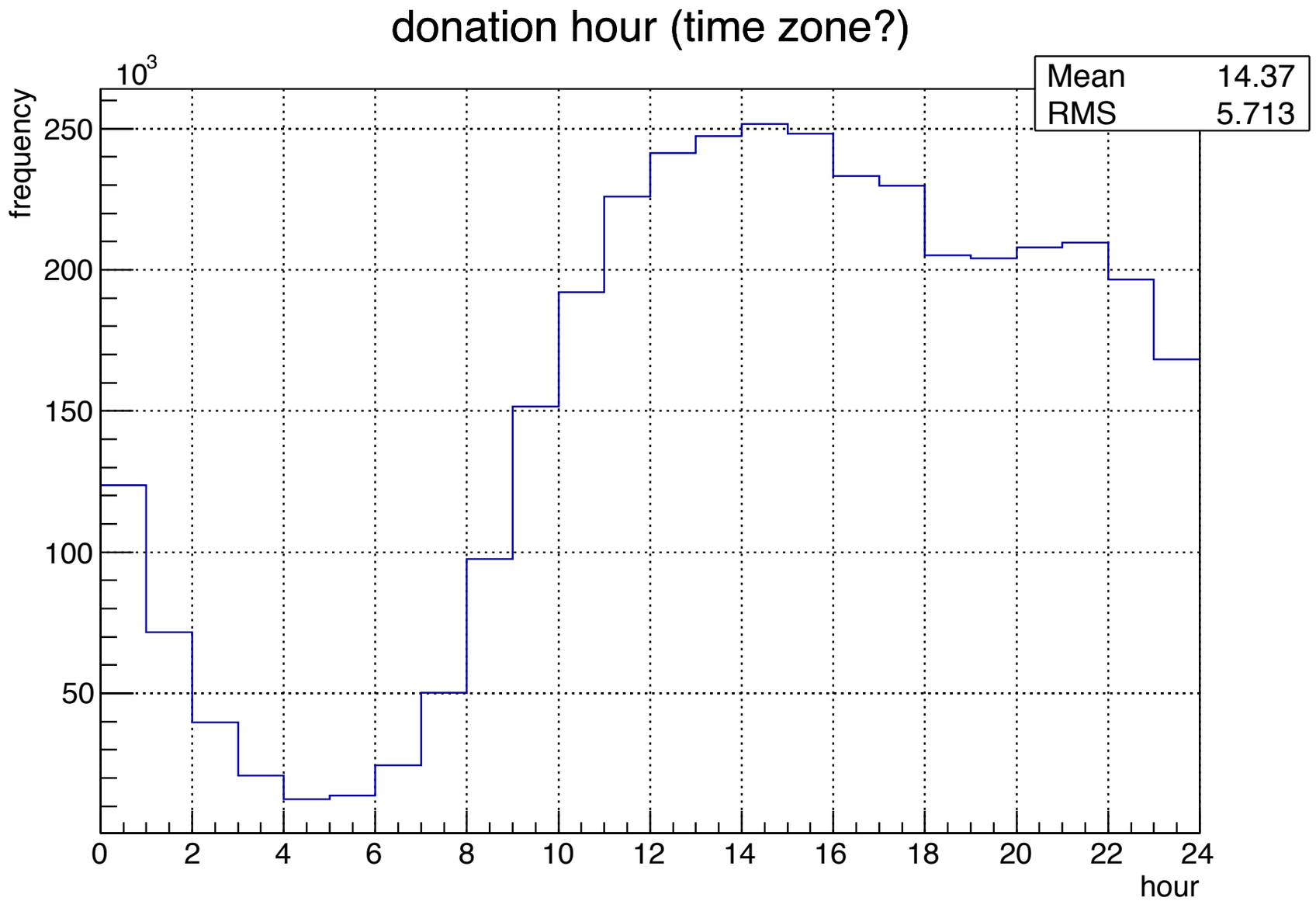
DSaPP activity

Joe Grange

Section 1: exploratory analysis

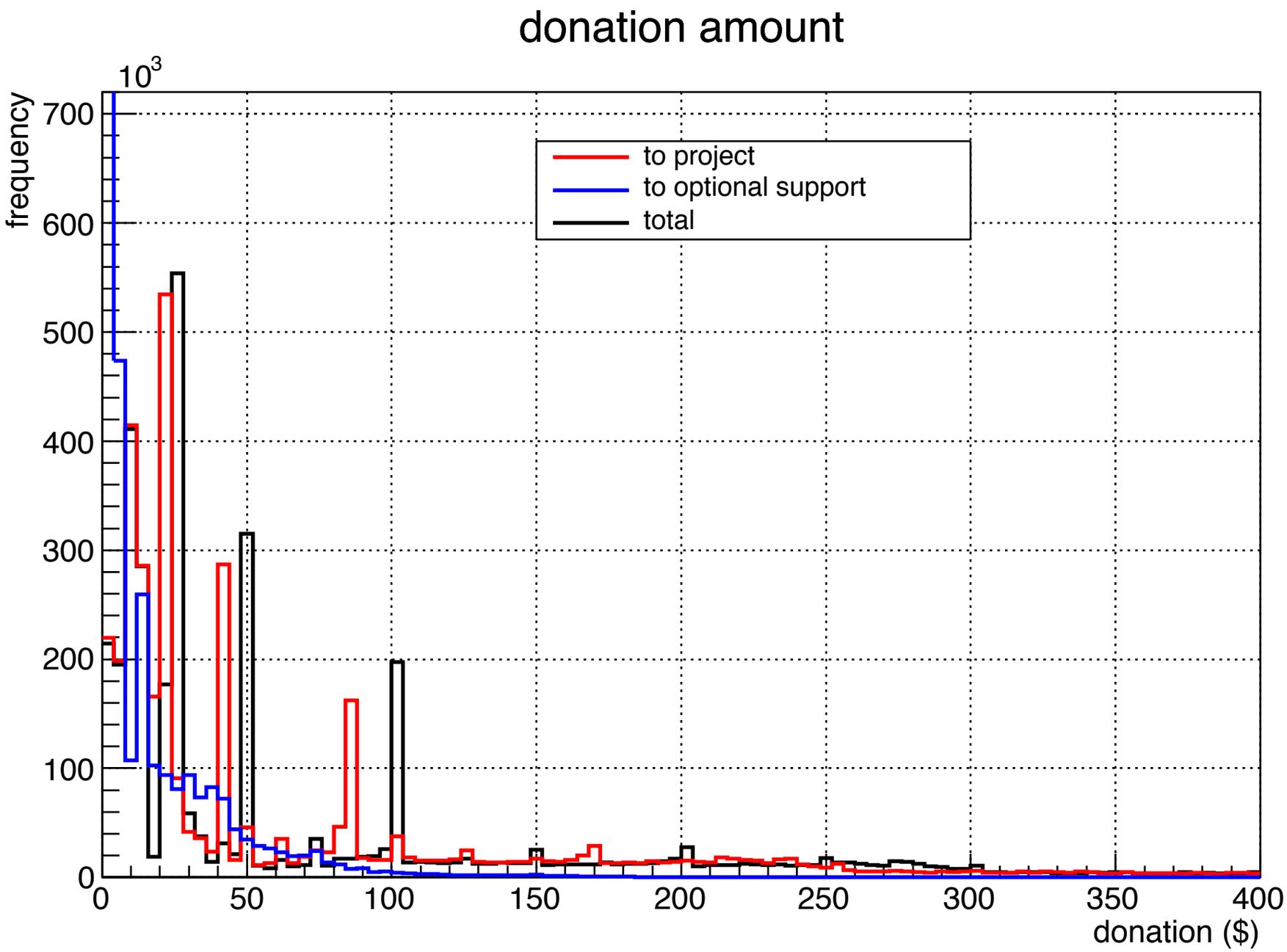
For fun: donation hour

- Surprising that roughly half as many donations at midnight compared to peak donation hours (though I'm not sure if time zone reflects the donor time zone or local CPU time)



Donations

- Multiples of $25 for the total donation are common

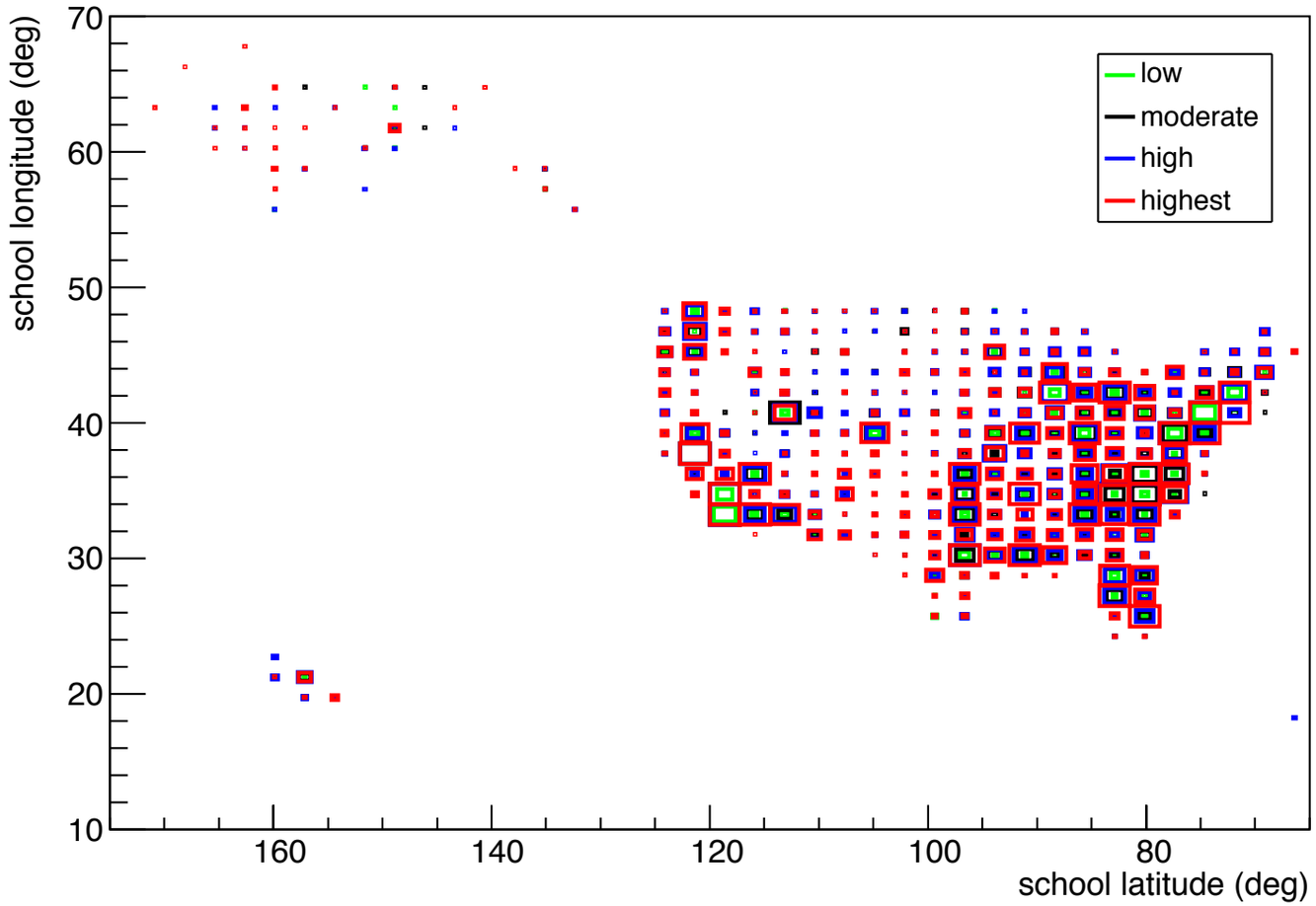


Project poverty by region

- Size of box proportional to entries in the bin divided by total for each distribution
  - highest poverty group more evenly distributed across the country compared to the others

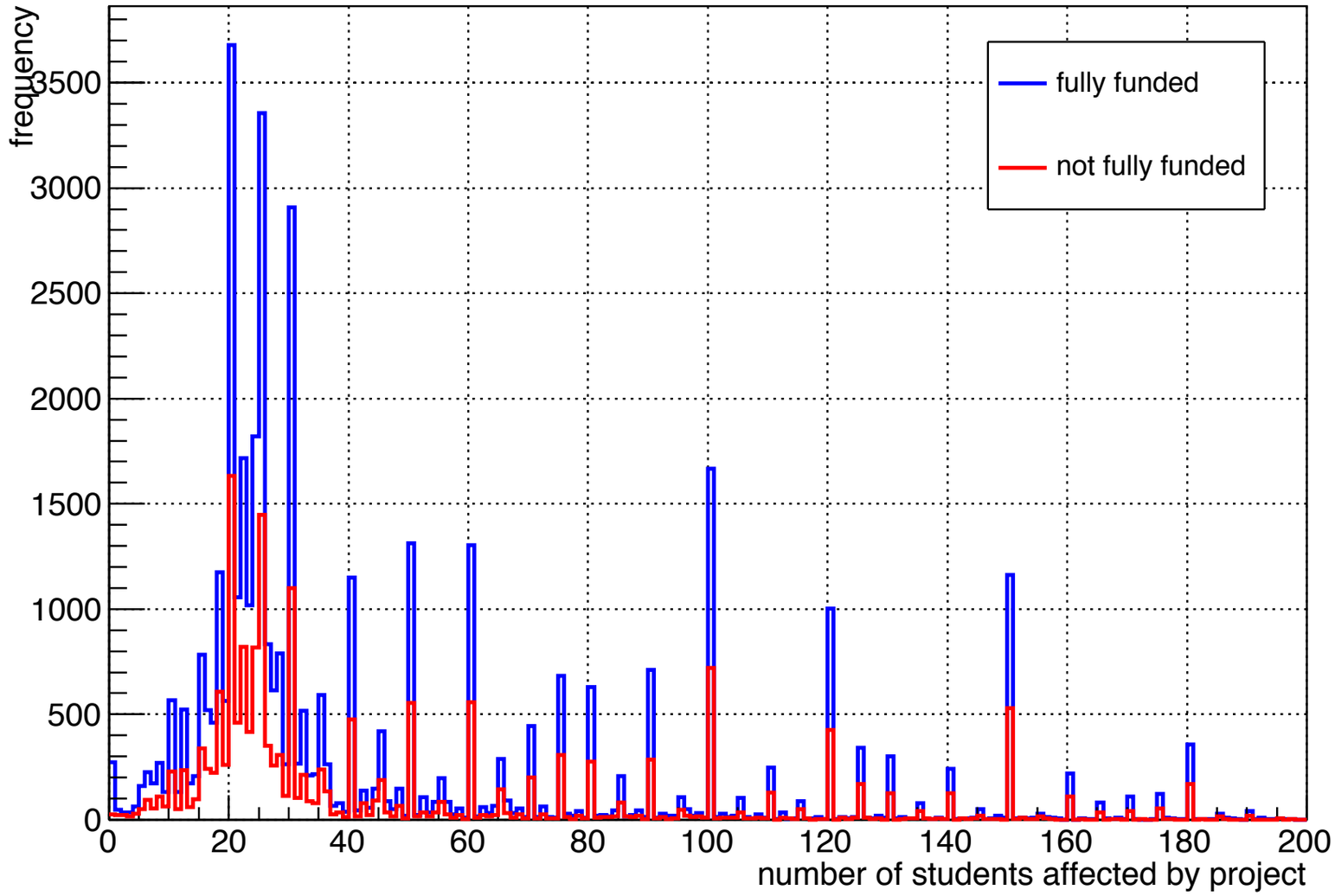
- A US population map probably looks very similar to this

project poverty distribution (shape only)



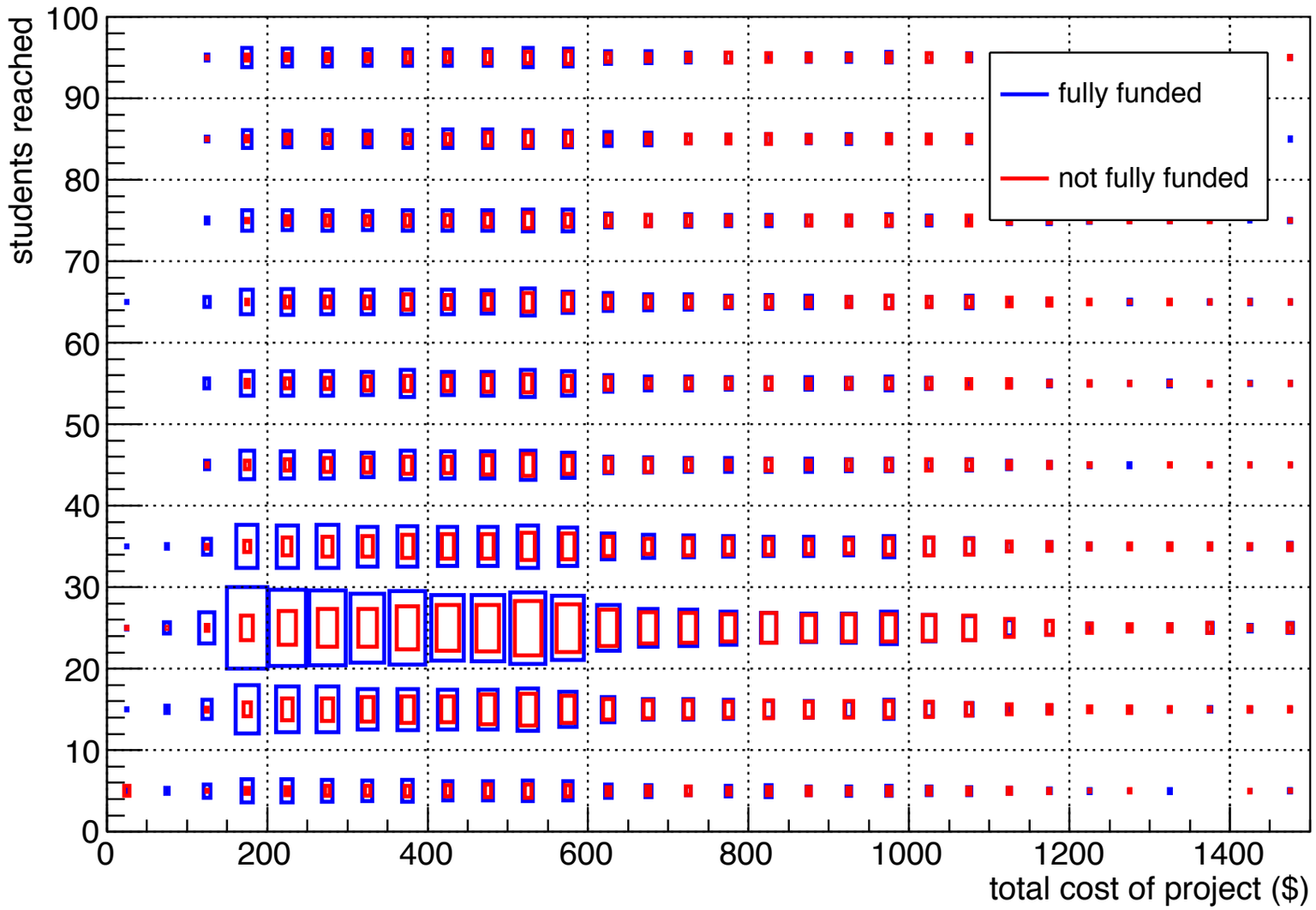
Number of students reached and funding level

- Similar shape to both distributions - no clear funding bias depending on number of students reached
- "Rounding effect" - clear peaks at multiples of 50, 10, 5



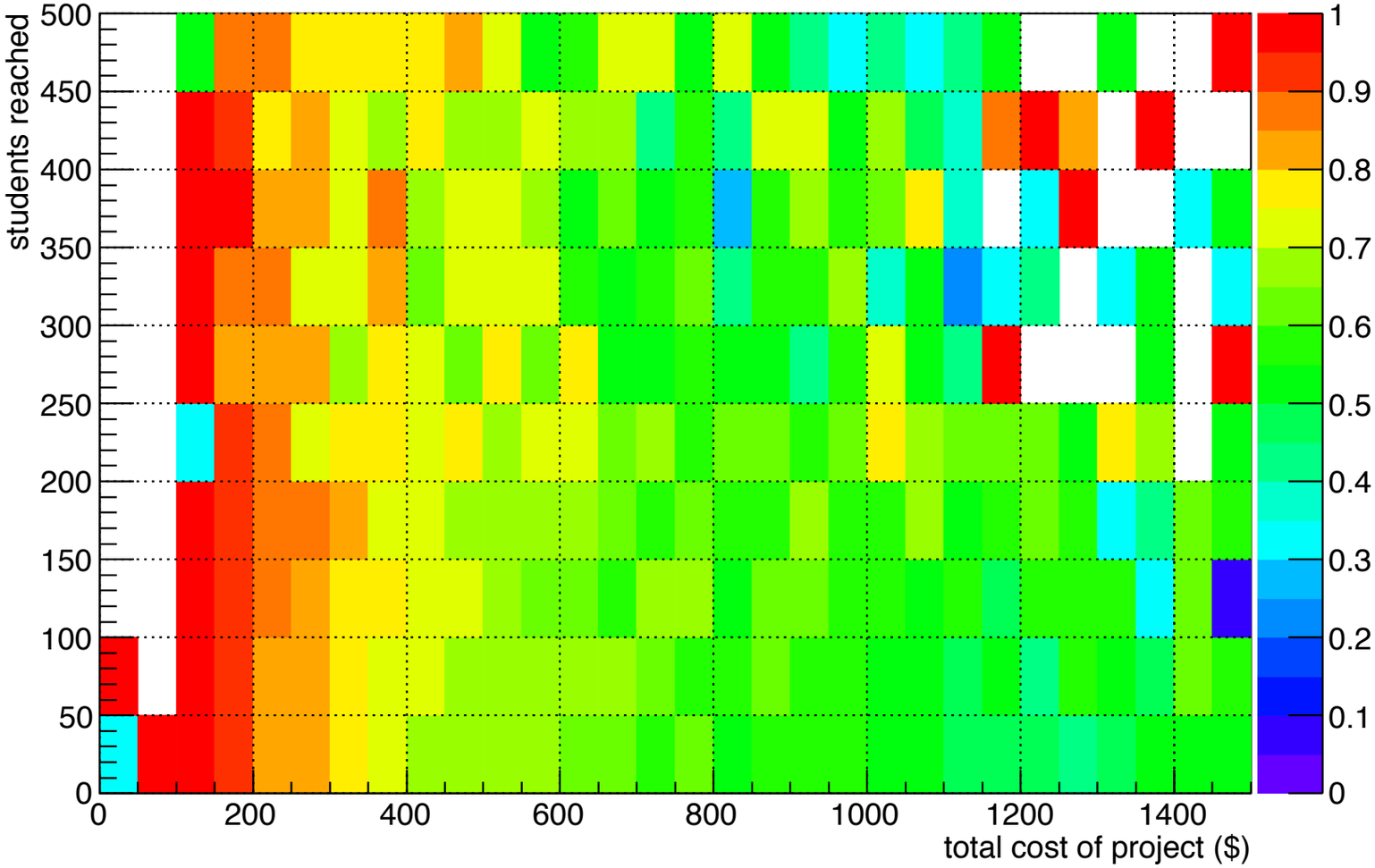
Students reached vs cost vs funding

- "Sweet spot" for fully funded projects around typical class size (20-30) and < $600.
- Not-fully funded projects tend to have higher total costs
  - not surprising



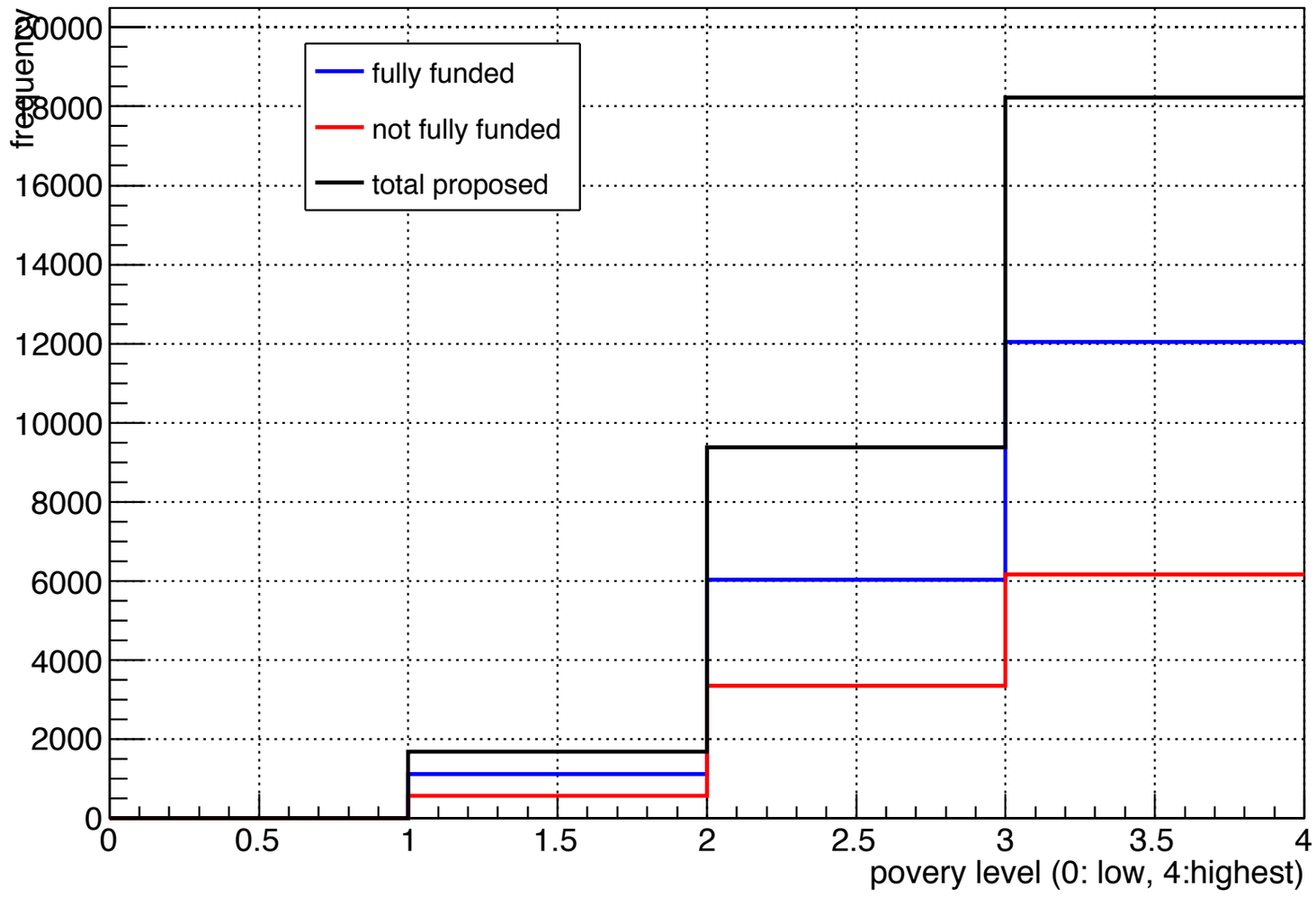
Let's look at the ratio

- Fully funded projects / total by project cost and students reached.
- If the project cost is < $300 or so, very high chance of being fully funded
  - were these funded through donations only or are there more agents involved? If there were some funding committee, I'd be suspicious of such a dependence…
- No dependence here on the number of students reached



Poverty level by funding success

- Roughly equal proportions funded in each poverty group



Some tables

Field	% of total
Exciting project	6%
At least 1 teacher referred donor	20%
Fully funded	70%
At least 1 green donation	61%
Great chat	30%

Grades affected by project	% of total
PreK-2	37%
3-5	31%
6-8	17%
9-12	15%

Teacher gender statistics

- I find this pretty interesting: projects proposed by male teachers were fully funded at a much lower rate than female-teacher-led projects (factor 6 different).

Gender	Total	% fully funded	% exciting
Female	60379	60%	5%
Male	8788	9%	0.8%

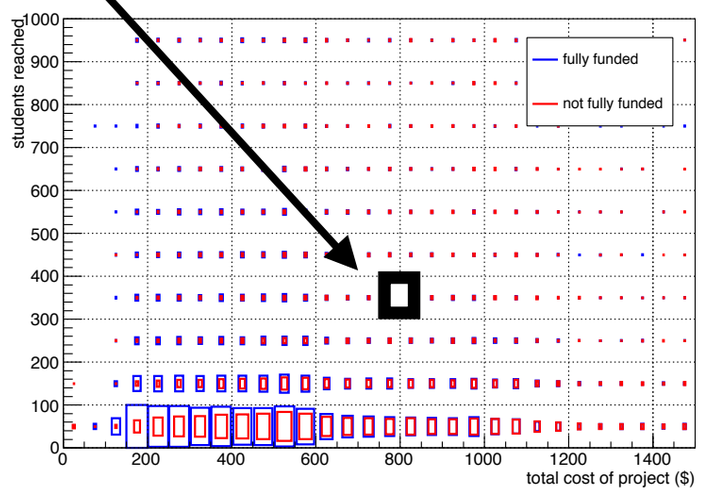
- I decimated the data set to produce this table, only analyzing ~1/10 of the total data. As long as there's no bias in the first 10% as ordered in outcomes.csv, conclusions drawn here are safe.

Section 2: Data story

To be honest I'm not sure what you folks are
looking for
in this section.
It seems to me the power of this
(and similar) data sets is solely
in the large statistics,
and so it doesn't seem helpful to me
to look too closely at individual cases.

Projectid e44a4b2d402e71746ed5853a955778ef

- Male teacher
- Highest poverty class
- Large reach: 400 students affected
- Low cost ($800) for so many students reached
- Fully funded, “exciting” project
- One of few fully funded and exciting projects led by a male teacher



Section 3: Q's for project partner

General and partner questions

- The low proposal and success rate for male teacher projects seems interesting.
  - General question: does the low male proposal rate just reflect the population of male vs female teachers? If so, nothing much to do about that. Partner question: if not, why weren't more male teachers engaged in project proposal? Could imagine some additional studies possible.
  - Partner question: why is the funding and "exciting" rate so much lower for males than female teacher projects? Perhaps the answer could already be gleaned from this large data set. I would look for more correlations between male and female led projects to see if the apparent bias could be explained by something more normal, for example the projects proposed by males were simply more expensive/less viable than those proposed by females.

- One field I think could've helped interpret this data is some measure of impact on the students' lives/schooling. We have number of students reached, but does the project have high impact (e.g. better nutrition at school) or low impact (e.g. 20% more pencils in the classroom)

General and partner questions

- One field I think could've helped interpret this data is some measure of impact on the students' lives/schooling. We have number of students reached, but does the project have high impact (e.g. better nutrition at school) or low impact (e.g. 20% more pencils in the classroom).
  - this information is in the essays but a more coarse evaluation would be helpful