

Fall 2025 CAMCOS Wildfire project

Meeting with Wilkin Lab

2025-09-11

Data Issues

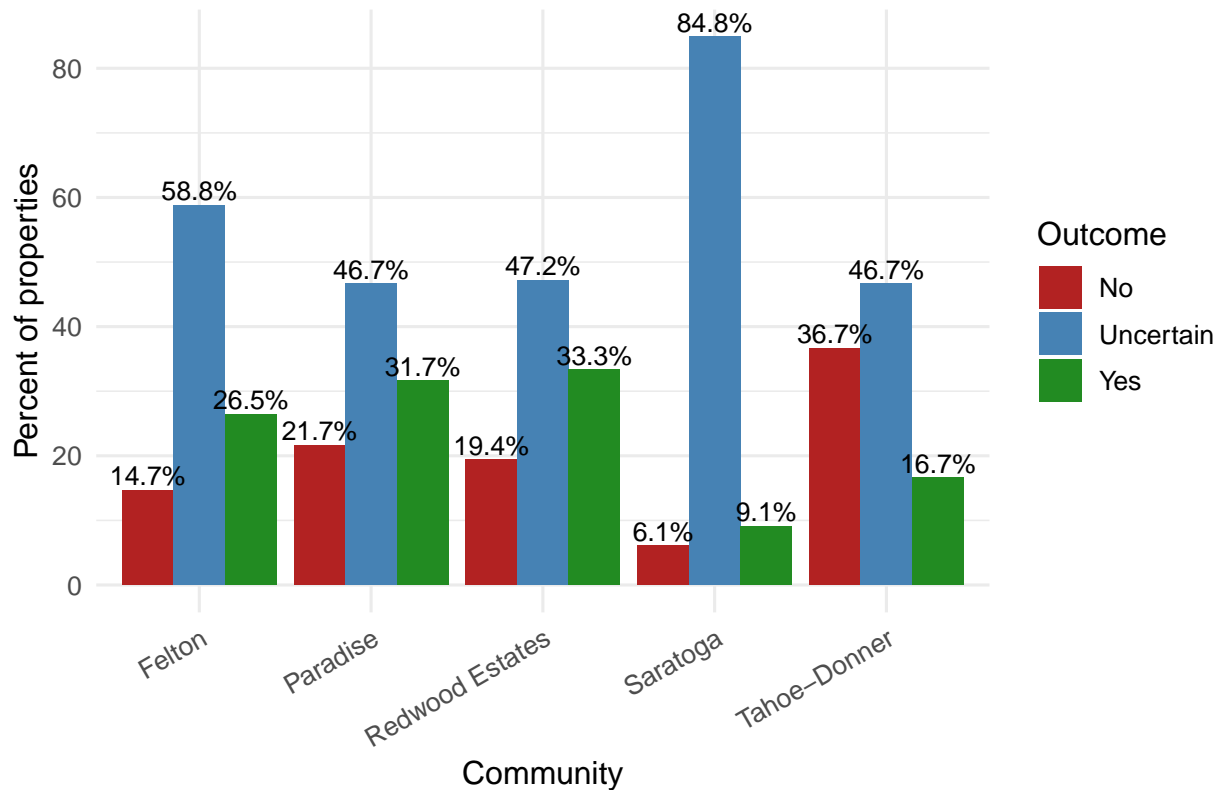
We took a closer look at the data and noticed a couple of issues that were mentioned in the email from Martina to Kate and Basanth on 9/2/25. We should talk about these and resolve questions (if they haven't already been addressed.)

The questions were: - What is Parcel ID? An alternative enumeration of properties? - We noticed that several properties (the ones that appear in the bottom rows (176 - 222) of the original file (tab: "Field Data") have duplicates among them. That is, the same "UID" is repeated across several rows. The results in these rows differ (different built years, different roof types etc.) For example, property "727e093140" appears in rows 200 - 204 with two shingle roofs and three metal roofs. These rows all have no "year surveyed" in column F. What's up with these properties? Are these consecutive visits to the same property over time? Do you want us to use these data? Or only the most recent observation for each house? Using repeated observations on the same house would make the model to be fitted quite a bit more complicated. - There is information on the "most recent Cal Fire" inspection. Is there also information on when that most recent inspection took place? - Do we know how the slope value was calculated? There seems to be one property with a slope of ~55degrees. That's unbelievably steep. - Some columns (for instance "attic Vents gap size" and "attic Vent materia;") have the entry "unknown" in all rows. I have some memory of you talking about measuring vent gap sizes with a golf tee. That does not make sense if the data set has no known measurements. Are we sure that there is no information about these two variables at all?

Some Descriptive Statistics

The CAMCOS team has spent a week summarizing parts of the data to look for patterns that can help us find a good model to fit later on. Here are some of the graphs that have found interesting so far.

CAL FIRE Inspection Outcomes by Community



Across all communities, a majority of properties fall into the “Uncertain” category, making it difficult to draw strong conclusions on compliance.

- Saratoga leads with about 85% uncertain outcomes and very few clear passes or fails, suggesting incomplete or inconsistent reporting.
- Felton follows with nearly 59% uncertain. Paradise and Redwood Estates each have about one-third of homes passing, but nearly half uncertain.
- By contrast, Tahoe-Donner stands out with the highest failure rate (37%) and a low pass rate (17%), showing more properties in need of fire safety improvements.

Overall, the dominance of uncertain results limits clarity, but Tahoe-Donner reveals the most definite challenges.