Part 1: What information did farmers give us?

Edmund, Debbie, Andrew, Lindsay, and Silas gave us raw yield data. What did they think was relevant to record? I coded it in this spreadsheet.

I split it up into topics:

Identity - Crop, Variety, Latin name
Area/Spacing - number of plants, actual row feet, actual bed feet, plant spacing, area in square/bed feet, number of flats/cells sown in greenhouse, rows per bed
Planting - Date direct sown/seeded/transplanted
Environmental Factors - Year, soil type, irrigation, fertility interventions, pest interventions (often covered under notes")
Inputs - how much seed and/or soil was used during planting
Pollination - Hand-pollination date, pollination interventions (screen-house, detasseling)
Harvest - First edible harvest, first seed harvest, last seed harvest
Yield - Total yield, yield for sale
Info - Isolation distance, germination standard, whether grown for production/varietal improvement/etc.
Inputs - Amount of seed and potting soil used.
Financial - Amount contracted, price per pound, financial rating
Germination - Testing date, percentage

What was most important/consistent in folks' reporting?

Identity. Everyone identified their crop and variety.
Yield. We asked for yield and everyone gave that, either the amount sold to a seed company or the total harvest weight.
Notes. [Analyze what people include in notes -- mostly environmental and harvest info???]

People also included:

Area/spacing. The way in which it reported VARIED WIDELY, but everyone reported this in some fashion. Everyone gave enough information to calculate yield per acre.
Financial. Price per lb was recorded by most, and if they did, they also included the company it was being grown for.

Things people were least likely to record routinely, but to lump in with the notes if something unusual happened:
Environmental. This was most commonly included in the notes.
Harvest and planting dates. Do we need to know this?
Pollination.

Germination.
Inputs.
Info. We do not need to collect data on isolation distances, germination standards, etc. from farmers.

Part 2: Proposed collection form

A. Generic Form

I made a proposed collection form using JotForm. https://form.jotform.com/221493987643065

This form could be filled out by any farmer for any data they collected during any year. We could widely circulate it and see how much data we receive this way.

B. Customized Two-Part Form

I think that the best way to get this information for 2022 seed crops would actually be to have two different forms, an early-season form and a post-harvest form, and send them out to all the commercial seed growers that we know.

The early-season form would ask for:

- Farmer name and contact information
- Farm location
- Crops and varieties being grown this year
- Planting dates (or planned planting dates)
- Area planted (or planned area to plant)

The post-harvest form would ask for the yield as well as all the questions from the Generic Form under “Additional Information.”

I suggest this for two reasons:

- We could make customized forms so that farmers only have to select their name and don’t even get prompted for their contact information, etc. I believe I can make customized links for farmers where we pre-fill data about themselves AND their crops, which will increase responsiveness.
- The first form is relatively easy and may get more farmers to buy into the idea. Then we can give them reminders if they don’t fill out the more difficult second half, and likely get more data that way.

If we wanted to do this, the time for the early-season form (for farmers in my area) is definitely June, when they get a little downtime between planting and harvesting.