

Data Assumptions – Ella Rises (INTEX Fall 2025)

Use the following assumptions as you work through the normalization process.

Participants (People)

- Each row may reference a participant, but the same participant may appear in multiple rows.
- ParticipantEmail uniquely identifies a person.
- A participant's personal information does not change based on the event.

Events vs Event Occurrences

- EventName identifies a general type of event.
- EventName, EventType, EventDescription, EventRecurrencePattern, and EventDefaultCapacity describe the event template.
- EventDateTimeStart, EventDateTimeEnd, EventLocation, EventCapacity, and EventRegistrationDeadline describe a specific scheduled occurrence.
- A specific event occurrence is uniquely identified by EventName + EventDateTimeStart.

Registrations

- A registration represents a participant signing up for a specific event occurrence.
- A participant may register for multiple event occurrences.
- A participant may only register once per event occurrence.

Surveys

- A participant may submit at most one survey per event occurrence.
- Survey columns describe feedback from a participant about a specific occurrence.
- Not all registrations have surveys.

Milestones

- MilestoneTitles may contain multiple milestone titles separated by semicolons.
- MilestoneDates contains corresponding dates, in the same order.
- Each milestone represents a participant achieving something.
- Each participant may have zero or more milestones.

- Each milestone consists of the participant, the title, and the date achieved.

Donations

- DonationHistory contains donations separated by semicolons (YYYY-MM-DD:\$Amount).
- Each donation represents a single donation by the participant.
- TotalDonations is the sum of DonationHistory.
- A participant may have zero or more donations.

General Spreadsheet Rules

- A single row may contain participant, event, registration, survey, milestone, and donation data.
- If data is blank, create an entity only if enough identifying information exists.
- The spreadsheet is an unnormalized operational export.
- Duplicate data refers to the same real-world entity.

Normalization Expectations

- Resolve any non-atomic data and repeating groups.
- Identify entities and primary keys based on all these assumptions.
- Separate event templates from occurrences.
- Separate registrations from surveys.
- Move multi-valued attributes into separate tables.
- Eliminate partial and transitive dependencies.
- Create surrogate keys only after 3NF.