

| | |
|------------------|-----------------------|
| From: | Lead Engineer |
| Received: | Today |
| Subject: | Engineering questions |

Hello,

Sure – let's see if I can answer some of your questions.

- Each of those datasets comes from a different source; historically these were built independently by their teams. I can't speak to any standardisation practices or validation rules. I'm not surprised to hear about missing or non-standard data; certainly our systems go down sometimes.
- I don't think we do any cleaning, transformation or anonymisation as part of the pipelines. The only thing we do is calculate the resolution time for tickets in ETL.
- Each dataset gets updated daily.
- Missing or irregular data can occur if an upstream system experiences downtime. We don't automatically backfill gaps unless there's a major incident.
- Location for all fields is based on user billing data as far as I know.

Have a good afternoon,

Nicole
(Lead Engineer)