

Demo

Use case

- Real use case simplified
- A tracker tool sending user information in different requests

Demo

Use case

- 2 different streams
- Events are generated at the same time
- They arrive in random order and with delay
- Stream 1: User data with language
- Stream 2: User data with country

Demo

How it works

- 2 source tables: one per stream
- 3 MV:
 - 1 MV to join forward
 - 1 MV to join backward
 - 1 MV to redo forward joins
- 1 final table: the final merged entities

Demo

Actual demo

```
(ee) → multijoin_stream_demo ./execute.sh
cleaning data
** Data Source 'ds_visits' truncated
** Data Source 'ds_origins' truncated
** Data Source 'ds_visits_origins' truncated
** Data Source 'ds_origins_stats' truncated
-> Inserting a visit for user-00 on 2021/12/03 17:42:27

-----
| visit_date          | fingerprint | language | country |
-----
| 2021-12-03 17:42:27 | user-00     | es-ES    |         |
-----

-> Inserting now the origin for user-00 that arrives late 2021/12/03 17:42:26

-----
| visit_date          | fingerprint | language | country |
-----
| 2021-12-03 17:42:27 | user-00     | es-ES    | ES       |
-----
* The country is updated with the old information

-> Inserting now a new origin for user-01

-----
| visit_date          | fingerprint | language | country |
-----
| 2021-12-03 17:42:27 | user-00     | es-ES    | ES       |
-----
* It does not add any new visits because the actual visit didn't arrive

-> And now the visit for user-01 that arrives after the origin we just inserted

-----
| visit_date          | fingerprint | language | country |
-----
| 2021-12-03 17:42:27 | user-00     | es-ES    | ES       |
| 2021-12-03 17:42:43 | user-01     | en-US    | UK       |
-----
* It shows two visits with the right country/locale
(ee) → multijoin_stream_demo
```

Demo

How we usually do it

- We try to not store final entities but derivative datasets from streams
- origin_stats example

Demo

Actual demo

```
← → ↻ 🔒 api.tinybird.co/v0/pipes/origin_stats.json

{
  - meta: [
    - {
      name: "country",
      type: "LowCardinality(String)"
    },
    - {
      name: "visits",
      type: "Int64"
    }
  ],
  - data: [
    - {
      country: "UK",
      visits: 1
    },
    - {
      country: "ES",
      visits: 1
    }
  ],
  rows: 2,
  - statistics: {
    elapsed: 0.000352276,
    rows_read: 2,
    bytes_read: 49
  }
}
```