

ML Feature Table Computation

Problem

As a data scientist, I want timeseries data with the total number of transactions a user made in the last 7 days from a given date.

The timeseries should have the following dimensions:

- transaction_id
- user_id

Proposed Solution

Prepare a SQL query to calculate required KPI using the data from the `transactions` table:

```
CREATE TABLE transactions
(
  transaction_id      UUID,
  date                DATE,
  user_id             UUID,
  is_blocked          BOOL,
  transaction_amount  INTEGER,
  transaction_category_id INTEGER
);
```

The desired table would look as following.

transaction_id	user_id	date	# Transaction within previous 7 days
ef05-4247	becf-457e	2020-01-01	0
c8d1-40ca	becf-457e	2020-01-05	1
fc2b-4b36	becf-457e	2020-01-07	2
3725-48c4	becf-457e	2020-01-15	0
5f2a-47c2	becf-457e	2020-01-16	1
7541-412c	5728-4f1c	2020-01-01	0
3deb-47d7	5728-4f1c	2020-01-12	0

Acceptance Criteria

- Postgres compatible [ANSI compliant](#) SQL query is designed.
- Steps of the query planner are described to explain what a database engine considers for effectively query execution.