```
-- Task 1: Bot vs. Human Changes
SELECT
    SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.BOT = TRUE) THEN 1
ELSE 0 END)) BOT_CHANGES,
    SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.BOT = FALSE) THEN 1
ELSE 0 END)) NONBOT_OR_HUMAN_CHANGES,
    COUNT (*) AS TOTAL_COUNT
    FROM  WIKIMEDIA_RECENTCHANGES
    GROUP BY 1
    EMIT CHANGES;
```

KSQLDB ›

default

Editor   Flow   Streams   Tables   Running queries

```
1    SELECT
2        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.BOT = true) THEN 1 ELSE 0 END)) BOT_CHANGES,
3        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.BOT = false) THEN 1 ELSE 0 END)) NONBOT_OR_HUMAN_CHANGES,
4        COUNT (*) AS TOTAL_COUNT
5        FROM  WIKIMEDIA_RECENTCHANGES
6        GROUP BY 1
7        EMIT CHANGES;
8
```

● Add query properties   Running...   Run   Stop

Data structure
STREAM

Total messages
24884

Messages/sec
51.21

Total message bytes
27582125

Message fields

• BOT_CHANGES

BOT_CHANGES	NONBOT_OR_HUM...	TOTAL_COUNT
32	95	127
18	66	84

```
-- Task 2: Bot vs. Human Changes (with 2 minute window)
CREATE TABLE COUNT_TUMBLING_WINDOW WITH
(KAFKA_TOPIC='COUNT_TUMBLING_WINDOW', PARTITIONS=1, REPLICAS=1) AS
SELECT
        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.BOT = TRUE) THEN 1
ELSE 0 END)) BOT_COUNT,
        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.BOT = FALSE) THEN 1
ELSE 0 END)) HUMAN_COUNT,
        COUNT (*) AS TOTAL_COUNT,
    TIMESTAMPTOSTRING(WIKIMEDIA_RECENTCHANGES.WINDOWSTART, 'yyyy-MM-
dd HH:mm:ss', 'UTC') AS WINDOW_START,
    TIMESTAMPTOSTRING(WIKIMEDIA_RECENTCHANGES.WINDOWSTOP, 'yyyy-MM-
dd HH:mm:ss', 'UTC') AS WINDOW_STOP
        FROM  WIKIMEDIA_RECENTCHANGES
        WINDOW TUMBLING (SIZE 60 SECONDS)
        GROUP BY TYPE
        EMIT CHANGES;
-- Task 2: Result
SELECT WINDOW_START, WINDOW_STOP, BOT_COUNT FROM
COUNT_TUMBLING_WINDOW EMIT CHANGES;
```

KSQLDB ›

default

Editor   Flow   Streams   Tables   Running queries

```
1    SELECT WINDOW_START,  WINDOW_STOP, BOT_COUNT FROM COUNT_TUMBLING_WINDOW EMIT CHANGES;
```

● **Add query properties**

Running...   Run   **Stop**

Data structure
TABLE

Total messages
37

Messages/sec
0.37

Total message bytes
5020

▶ ❚❚   🔍 Filter by keyword

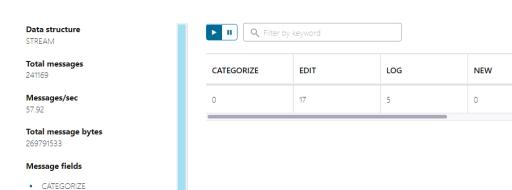| WINDOW_START	WINDOW_STOP	BOT_COUNT ^
2022-02-27 20:28:00	2022-02-27 20:29:00	337

```
-- Task 3: The Type field indicates the type of the change. What is
the distribution of Type over Windows of say 60 seconds?
SELECT
        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'categorieze')
THEN 1 ELSE 0 END)) CATEGORIZE,
        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'edit') THEN 1
ELSE 0 END)) EDIT,
        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'log') THEN 1
ELSE 0 END)) LOG,
        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'new') THEN 1
ELSE 0 END)) NEW,
        COUNT (*) AS TOTAL_COUNT,
    TIMESTAMPTOSTRING(WIKIMEDIA_RECENTCHANGES.WINDOWSTART, 'yyyy-MM-
dd HH:mm:ss', 'UTC') AS WINDOW_START,
    TIMESTAMPTOSTRING(WIKIMEDIA_RECENTCHANGES.WINDOWEND, 'yyyy-MM-dd
HH:mm:ss', 'UTC') AS WINDOW_STOP
        FROM  WIKIMEDIA_RECENTCHANGES
        WINDOW TUMBLING (SIZE 60 SECONDS)
        GROUP BY 1
        EMIT CHANGES;
```

```
1    SELECT
2        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'categorieze') THEN 1 ELSE 0 END)) CATEGORIZE,
3        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'edit') THEN 1 ELSE 0 END)) EDIT,
4        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'log') THEN 1 ELSE 0 END)) LOG,
5        SUM((CASE WHEN (WIKIMEDIA_RECENTCHANGES.TYPE = 'new') THEN 1 ELSE 0 END)) NEW,
6        COUNT (*) AS TOTAL_COUNT,
7        TIMESTAMPTOSTRING(WIKIMEDIA_RECENTCHANGES.WINDOWSTART, 'yyyy-MM-dd HH:mm:ss', 'UTC') AS WINDOW_START,
8        TIMESTAMPTOSTRING(WIKIMEDIA_RECENTCHANGES.WINDOWEND, 'yyyy-MM-dd HH:mm:ss', 'UTC') AS WINDOW_STOP
9        FROM  WIKIMEDIA_RECENTCHANGES
```

● Add query properties   Running... [Run] [Stop]

Data structure
STREAM

Total messages
241169

Messages/sec
57.92

Total message bytes
269791533

Message fields

- CATEGORIZE

- EDIT

CATEGORIZE	EDIT	LOG	NEW	TOTAL_C
0	17	5	0	24