*-- 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;

Graphical user interface, text, application

Description automatically generated

*-- 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;

Graphical user interface, text, application, email

Description automatically generated

*-- 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;

Graphical user interface

Description automatically generated