create table algorithmisation.person\_cluster\_all as select W.\*, CASE WHEN(Day\_of\_Week=6) THEN -1 WHEN(Day\_of\_Week=7) THEN -1 WHEN(call\_time>000000 AND call\_time<070000) THEN -1 WHEN(call\_time>090000 AND call\_time<170000 AND Day\_of\_Week!=6 AND Day\_of\_Week!=7) THEN 1 WHEN(call\_time>190000) THEN -1 ELSE 0 END AS dummy from (select A.caller\_msisdn, A.call\_date, A.call\_time, A.ms\_location, B.cluster\_id, B.centroid\_lat, B.centroid\_long, from\_unixtime(unix\_timestamp(call\_date,'yyyyMMdd'),'u') as Day\_of\_Week from (select S.caller\_msisdn, S.call\_date, S.call\_time, S.ms\_location, T.bts\_nodeb from algorithmisation.subset\_of\_cdr\_november S left join algorithmisation.antennas T on (S.ms\_location=T.id\_cellule)) A left join algorithmisation. clusters\_nov17 B on (A.bts\_nodeb=B.bts\_nodeb AND trim(A.caller\_msisdn)=trim(B.caller\_msisdn))) W;

create table algo\_july.person\_cluster\_all as select W.\*, CASE WHEN(Day\_of\_Week=6) THEN -1 WHEN(Day\_of\_Week=7) THEN -1 WHEN(call\_time>000000 AND call\_time<070000) THEN -1 WHEN(call\_time>090000 AND call\_time<170000 AND Day\_of\_Week!=6 AND Day\_of\_Week!=7) THEN 1 WHEN(call\_time>190000) THEN -1 ELSE 0 END AS dummy from (select A.caller\_msisdn, A.call\_date, A.call\_time, A.ms\_location, B.cluster\_id, B.centroid\_lat, B.centroid\_long, from\_unixtime(unix\_timestamp(call\_date,'yyyyMMdd'),'u') as Day\_of\_Week from (select S.caller\_msisdn, S.call\_date, S.call\_time, S.ms\_location, T.bts\_nodeb from algo\_july.subset\_of\_cdr\_july S left join algorithmisation.antennas T on (S.ms\_location=T.id\_cellule)) A left join algo\_july. clusters\_nov17 B on (A.bts\_nodeb=B.bts\_nodeb AND trim(A.caller\_msisdn)=trim(B.caller\_msisdn))) W;

create table algorithmisation.person\_cluster\_days as select A.\* from( select caller\_msisdn, cluster\_id, count(DISTINCT call\_date) as tot\_days from algorithmisation.person\_cluster\_all GROUP BY caller\_msisdn, cluster\_id)A WHERE tot\_days>3;

create table algo\_july.person\_cluster\_days as select A.\* from( select caller\_msisdn, cluster\_id, count(DISTINCT call\_date) as tot\_days from algo\_july.person\_cluster\_all GROUP BY caller\_msisdn, cluster\_id)A WHERE tot\_days>3;

create table algorithmisation.person\_cluster\_all2 as select A.caller\_msisdn, A.cluster\_id, A.centroid\_lat, A.centroid\_long,  (SUM(dummy) OVER (PARTITION BY caller\_msisdn, cluster\_id)) as tot\_cat from (select B.caller\_msisdn, B.call\_date, B.call\_time, B.ms\_location, C.cluster\_id, C.centroid\_lat, C.centroid\_long FROM algorithmisation.person\_cluster\_all B  inner join algorithmisation.person\_cluster\_all C on (B.caller\_msisdn=C.caller\_msisdn AND B.cluster\_id=C.cluster\_id)) A where tot\_cat<0 GROUP BY A.caller\_msisdn, A.cluster\_id, A.centroid\_lat, A.centroid\_long;

create table algo\_july.person\_cluster\_all2 as select A.caller\_msisdn, A.cluster\_id, A.centroid\_lat, A.centroid\_long,  SUM(dummy) OVER (PARTITION BY caller\_msisdn, cluster\_id) as tot\_cat from (select B.caller\_msisdn, B.call\_date, B.call\_time, B.ms\_location, C.cluster\_id, C.centroid\_lat, C.centroid\_long FROM algo\_july.person\_cluster\_all B  inner join algo\_july.person\_cluster\_all C on (B.caller\_msisdn=C.caller\_msisdn AND B.cluster\_id=C.cluster\_id)) A where tot\_cat<0 GROUP BY A.caller\_msisdn, A.cluster\_id, A.centroid\_lat, A.centroid\_long;