

study period = 03/25/2021 - 04/25/2021

covidtest

patient_id	covid_test_result	sample_collection_date
1001	Positive	2021-03-05
1001	Negative	2021-03-06
1001	Positive	2021-03-08

[filter]
covid_test_result = Positive
patient_id <> B00000000
sample_collection_date within period of study
[transpose]

patient_id	day_1 (2021-03-05)	day_2	day_3	day_4 (2021-03-08)	..	day_k (2021-04-03)
1001	1	0	0	1	0	0

hospcore - hospital admissions

patient_id	admit_date	diagnosis_code_1	..	diagnosis_code_20
1001	2021-03-08	U071
1001	2021-04-03	J128

[filter]
admit_date within period of study
patient_id <> B00000000
a Covid-related ICD code present in any diagnosis_code column
[transpose]

patient_id	day_1 (2021-03-05)	day_2	day_3	day_4 (2021-03-08)	..	day_k (2021-04-03)
1001	0	0	0	2	0	2

[merge]
combine rows with the same
patient_id by summing up the
columns

patient_id	day_1 (2021-03-05)	day_2	day_3	day_4 (2021-03-08)	..	day_k (2021-04-03)
1001	1	0	0	3	0	2

[transform]
4-step process
the goal is for each patient journey to
have at most one capture and one
recapture

h_2

patient_id	day_1 (2021-03-05)	day_2	day_3	day_4 (2021-03-08)	..	day_k (2021-04-03)
1001	1	0	0	2	0	0

obs_a

obs_b

[transform]
2s to 1s

h

patient_id	day_1 (2021-03-05)	day_2	day_3	day_4 (2021-03-08)	..	day_k (2021-04-03)
1001	1	0	0	1	0	0