**Cases Data Clearing Bad Data Guide**

Raw Data from Philippos: *CasesData.xls* (in folder Step1\_Manually)

* **Total** Raw cases before removing bad data: 473986
* **from** 07/03/2020 **to** 13/05/2022

**Bad data:**

1. Manually

**Folder:** *Step1\_Manually*

* Delete cases with missing Recovered Date.
* Delete cases that have Admission Date but no Discharged Date.
* Delete cases that died with no hospitalization.

File Result: *ClearedCasesData\_Manually.xls*

1. Script in Matlab

**Folder:** *Step2\_Script*

**File:** *ClearingCasesData.m*

Taking as input the result file from point 1 and also deleting the data that:

* Delete cases with Vaccination Date before the start date of vaccinations in Cyprus (27/12/2020).
* Delete cases that the date of Admission in hospital is before the date of the First Sampling.
* Delete cases that the Recovery date is before the First Sampling cases.
* Delete cases that the Discharged date is before the Admission date.

Result: *Results\_Cleared.mat* OR *ClearedCasesData.csv*

**Total cases after the deletion of bad data:** 469213

**Percentage of bad data: 1%**