README

****
About the NEWID
****

The original geographical identifiers in the DVN versions were scrambled in order to conform to public data dissemination protocol.

Please note that the "newid" for HH, KI, IT, and BAAC modules has slightly different composition. The scrambled "NEWID" still consistently represents unique geographical location (although not the real geo ID) of each case.

For non-mapping project, it is possible to track a given Village/ Tambon /Amphoe over time using the public data without the original codes:

For the Household Surveys 1997-2017, the NEWID is composed by:

1-2 digits = Changwat (province) code
3-4 digits = Amphoe (county) code
5-6 digits = Tambon (subcounty) code
7-8 digits = Village code
9-12 digits = Household code

Example...
072124510902
changwat: 07
ampho: 21
tambon: 24
village: 51
household: 0902

For the Key Informant Surveys 1997-2017, the NEWID is composed by:

1-2 digits = Changwat (province) code
3-4 digits = Amphoe (county) code
5-6 digits = Tambon (subcounty) code
7-8 digits = Village code

Example.
27324053
changwat: 27
ampho: 32
tambon: 40
village: 53

For the Institution Surveys 1997-2017, the NEWID is composed by:

1-2 digits = Changwat (province) code
3-4 digits = Amphoe (county) code
5-6 digits = Tambon (subcounty) code
7-8 digits = Village code
9-12 digits = Institution code

Example...
532938571812
changwat: 53
ampho: 29
tambon: 38
village: 57
institution: 1812

For the BAAC Surveys 1997 & 2000, the NEWID is composed by:

1-2 digits = Changwat (province) code
3-4 digits = Amphoe (county) code
5-6 digits = Tambon (subcounty) code
7-8 digits = Village code
9-11 digits = Group code

Example.

07212451008
changwat: 07
ampho: 21
tambon: 24
village: 51
group: 008

By extract the first 8 digits from the "NEWID", a user can obtain the village dummies.
In STATA, we can simply use the following commands to extract the substring:

```
gen villageid=substr(newid, 1, 8)
```