

For Species Distribution data (24239 observations):

Column Name	Description	Type	Mean	Median	Range	#Unique Values	# NAs
Country	Name of the country.	Categorical	/	/	/	47	0
GAUL_Admin1	UN first level administrative unit (2008).	Categorical	/	/	/	518	0
GAUL_Admin2	UN second level administrative unit (2008).	Categorical	/	/	/	1976	0
Lat	Latitude.	Numerical	1.199	4.788	-33.925 to 22.102	/	82
Long	Longitude.	Numerical	17.852	15.977	-24.975 to 50.773	/	82
YeStart	Year of start of the study.	Numerical	1976	1981	1898 to 2016	119	0
YeEnd	Year of end of the study.	Numerical	1989	2000	1900 to 2016	112	0
Complex¹	Name of the complex the observed mosquito belongs to.	Categorical	/	/	/	5	0
Species	Species name of the observed mosquito.	Categorical	/	/	/	20	0
Sub_Species	Sub-species name of the observed mosquito. ²	Categorical	/	/	/	4	17449
Multiple_methods	Whether or not the study used multiple (more than one) methods of identification.	Binary	/	/	/	2	0

¹ Detailed information regarding the Complexes, Species and Sub-species can be found in the file named “Species Dictionary”.

² If the species has no sub-species the entry is NA. If the species has sub-species but is not identified, the entry is “Unspecified Sub Species”.

For Species Richness and Malaria Cases data (301 observations):

Column Name	Description	Type	Mean	Median	Range	#Unique Values	# NAs
Country	Name of the country ³	Categorical	/	/	/	301	0
Year	Year in which the observation was made.	Numerical	2011	2010	2007-2016	10	0
Species Richness	Country level Species Richness ⁴ of Anophelines of that year.	Numerical	5.86	6	1-10	10	0
Incidence	Country level Malaria incidence (the number of malaria cases per 1000 people)	Numerical	235.59	243.04	0.05 - 552.03	/	0
Cases	The number of malaria cases that have been confirmed by examination.	Numerical	857162	239787	0 - 12293820	/	0
Rural %	Percentage of the country's total population living in rural areas.	Numerical	60.17	60.80	12.84 - 90.14	/	0
Rural Pop Growth %	Percentage annual change of a country's rural population.	Numerical	1.692	1.900	-2.530 - 7.090	/	0

³ One country (Mayotte) in the species distribution dataset was not included in this dataset and was thus removed. Inconsistencies in the country names have been fixed (species distribution dataset used as standard).

⁴ Species Richness was calculated as the number of different complexes (as species in the same complex are morphologically indistinguishable).