# Data dictionary

This data dictionary is a document that describes the basic organization of the tables in our project. This document contains a list of variables in the tables as well as the assigned variable names and a description of each variable. The data dictionary is used primarily for data analysis.

[Data dictionary 1](#_Toc531086000)

[1. Biome 2](#_Toc531086001)

[2. Continent 2](#_Toc531086002)

[3. Continent2 3](#_Toc531086003)

[4. Convex 4](#_Toc531086004)

[5. Convex\_with\_year 4](#_Toc531086005)

[6. Distribution 5](#_Toc531086006)

[7. Diversity 5](#_Toc531086007)

[8. Extinction 6](#_Toc531086008)

[9. Extinction\_with\_year 6](#_Toc531086009)

[10. Extinction\_in\_combined\_environments 6](#_Toc531086010)

[11. Lat 6](#_Toc531086011)

[12. Lon 6](#_Toc531086012)

[13. Niche\_breadth 7](#_Toc531086013)

[14. Origin 8](#_Toc531086014)

[15. Speciation\_Convex 8](#_Toc531086015)

[16. Speciation\_Convex\_E 8](#_Toc531086016)

[17. Speciation\_Extinction\_BIOME 8](#_Toc531086017)

[18. Speciation\_Extinction\_Continent 8](#_Toc531086018)

[19. Speciation\_Extinction\_Continent2 8](#_Toc531086019)

[20. Speciation\_Extinction\_Lat 8](#_Toc531086020)

[21. Speciation\_Extinction\_TT 8](#_Toc531086021)

[22. Speciation\_extinction\_in\_combined\_environments\_XXX 8](#_Toc531086022)

[23. Simulations 9](#_Toc531086023)

[24. Temperate\_Tropic 9](#_Toc531086024)

[25. Tree 9](#_Toc531086025)

## Biome

This table is used to store the distribution (number of cells) of each species (including the original and new species) within a biome. We have 11 different type of biomes. The original reference has 9 biomes. I split the desert in to 3 parts by the annual temperature. See figure below.



|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| BIOME | INT(11) | [1, 11] |
| COUNT\_BIOME | INT(11) | Number of the cells occupied by the species in year X |
| SPECIES\_ID | VARCHAR(45) | Species id within a simulation |
| YEAR | INT(11) | [100, 1200000] by =100 |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |

## Continent

This table is used to store the distribution (number of cells) of each species (including the original and new species) within a continent.

1: Africa 2: Asia 3: Australia 4: Europe 5: North America 6: South America

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| CONTINENT | INT(11) | [1, 6] |
| COUNT\_ CONTINENT | INT(11) | Number of the cells occupied by the species in year X |
| SPECIES\_ID | VARCHAR(45) | Species id within a simulation |
| YEAR | INT(11) | [100, 1200000] by =100 |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |

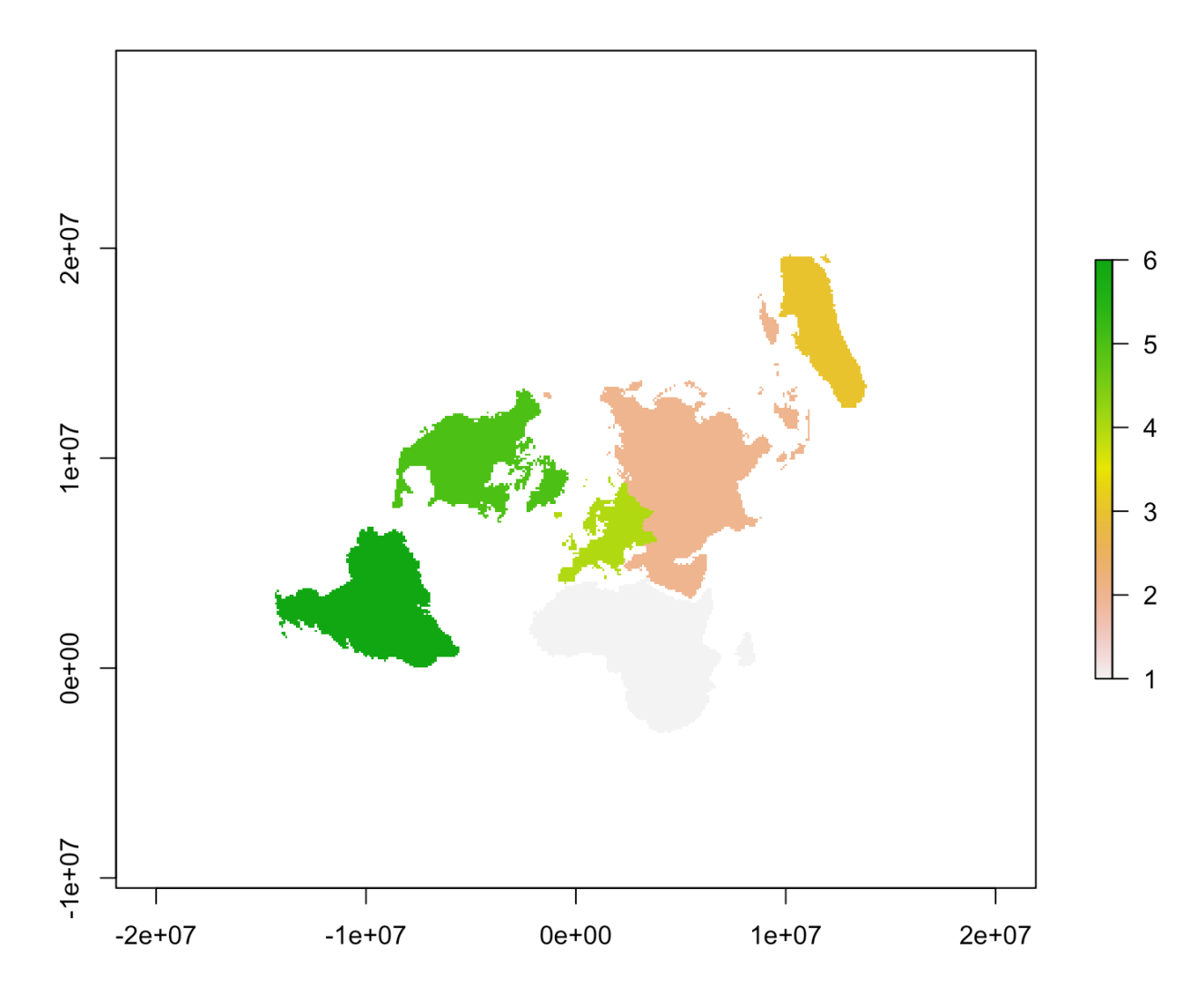
## Continent2

This table is used to store the distribution (number of cells) of each species (including the original and new species) within a continent groups.

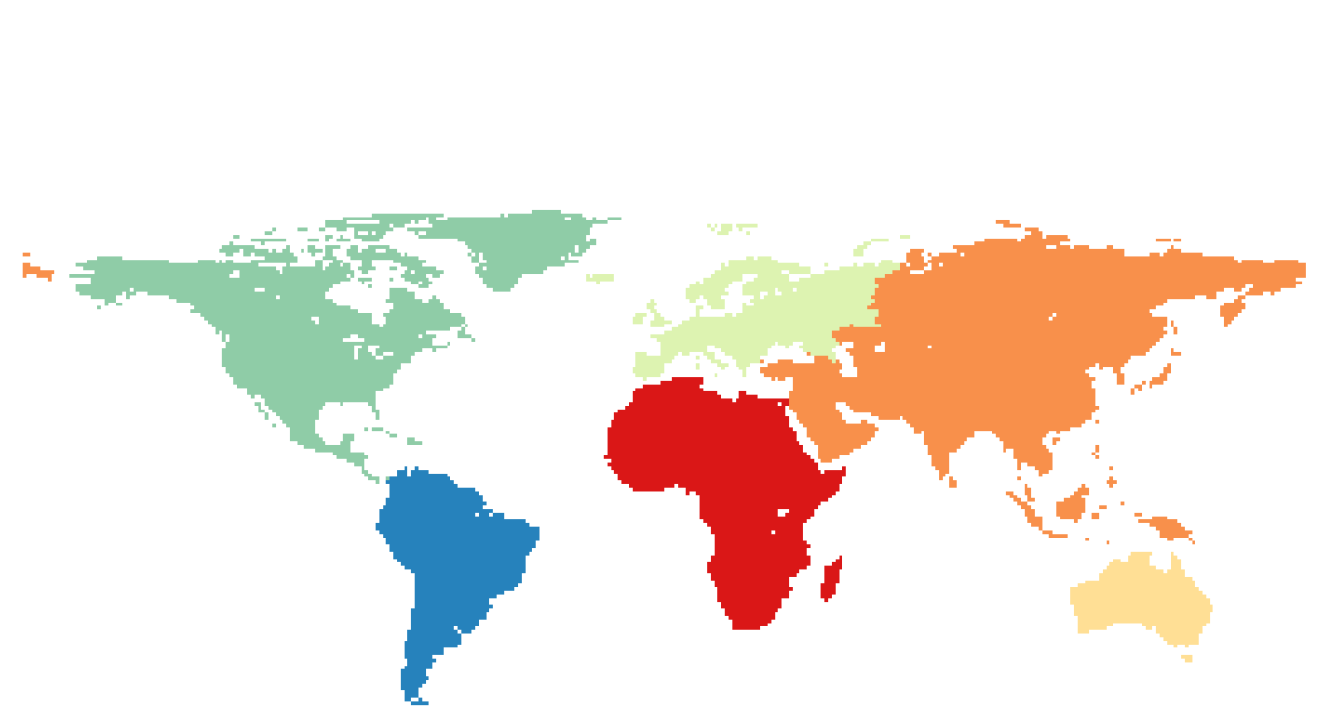
1: Africa and Europe 2: Asia and Australia 3: North and South America

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| CONTINENT | INT(11) | [1, 3] |
| COUNT\_ CONTINENT | INT(11) | Number of the cells occupied by the species in year X |
| SPECIES\_ID | VARCHAR(45) | Species id within a simulation |
| YEAR | INT(11) | [100, 1200000] by =100 |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |

**Note:** We used three different projections in simulation Round 10 (equal distance), 11 (equal area) and 12 (longlat).



Continents for R10



Continents for R11

## Convex

This table saved the pixels where the speciation happened. (Replaced by Table ‘Convex\_with\_year’)

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| x | FLOAT | Longitude |
| y | FLOAT | Latitude |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |
| v | INT(11) | 0: speciation area, 1-n: the ID of populations. |

## Convex\_with\_year

This table saved the pixels where the speciation happened.

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| x | INT(11) | X index in the map matrix |
| y | INT(11) | Y index in the map matrix |
| v | INT(11) | 0: speciation area, 1-n: the ID of populations. |
| isV | INT(11) | 1: is a vertex of the convex hull |
| year | INT(11) | Speciation year |
| species\_id | VARCHAR(45) | Species id within a simulation |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |
| lon | FLOAT | Longitude |
| lat | FLOAT | Latitude |

## Distribution

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| COUNT\_ DISTRIBUTION | INT(11) | Number of the cells occupied by the species in year X |
| SPECIES\_ID | VARCHAR(45) | Species id within a simulation |
| YEAR | INT(11) | [100, 1200000] by =100 |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |

## Diversity

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) |  |
| x | FLOAT |  |
| y | FLOAT |  |
| sp\_id | INT(11) |  |
| niche\_breadth | VARCHAR(45) |  |
| dispersal\_ability | VARCHAR(45) |  |
| v | INT(11) |  |

## Extinction

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) |  |
| x | FLOAT |  |
| y | FLOAT |  |
| sp\_id | INT(11) |  |
| niche\_breadth | VARCHAR(45) |  |
| dispersal\_ability | VARCHAR(45) |  |
| v | INT(11) |  |

## Extinction\_with\_year

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) |  |
| x | FLOAT |  |
| y | FLOAT |  |
| year | INT(11) |  |
| species\_id | VARCHAR(45) |  |
| sp\_id | INT(11) |  |
| niche\_breadth | VARCHAR(45) |  |
| dispersal\_ability | VARCHAR(45) |  |
| lon | FLOAT |  |
| lat | FLOAT |  |
| max\_prec | INT(11) |  |
| max\_temp | INT(11) |  |
| min\_temp | INT(11) |  |
| FN\_MIN\_TEMP\_MIN | INT(11) |  |
| FN\_MIN\_TEMP\_MAX | INT(11) |  |
| FN\_MAX\_TEMP\_MIN | INT(11) |  |
| FN\_MAX\_TEMP\_MAX | INT(11) |  |
| FN\_MAX\_PREC\_MIN | INT(11) |  |
| FN\_MAX\_PREC\_MAX | INT(11) |  |
| is\_last\_year | INT(11) |  |
| last\_appear\_year | INT(11) |  |

## Extinction\_in\_combined\_environments

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| sp\_id | INT(11) |  |
| niche\_breadth | VARCHAR(45) |  |
| dispersal\_ability | VARCHAR(45) |  |
| Band\_T\_MIN | INT(11) |  |
| Band\_T\_MAX | INT(11) |  |
| Band\_P\_MAX | INT(11) |  |
| Value | INT(11) |  |
| SE | VARCHAR(45) |  |
| ID | INT(11) |  |

## Lat

This table is used to store the distribution (number of cells) of each species (including the original and new species) within a latitude band(one degree).

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| LAT | FLOAT | [-55, 84] 139 |
| COUNT\_LAT | INT(11) | Number of the cells occupied by the species in year X |
| SPECIES\_ID | VARCHAR(45) | Species id within a simulation |
| YEAR | INT(11) | [100, 1200000] by =100 |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |

## Lon

This table is used to store the distribution (number of cells) of each species (including the original and new species) within a longitude band(one degree).

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| ID | INT(11) | identity |
| LON | FLOAT | [-180, 180] 360 |
| COUNT\_LON | INT(11) | Number of the cells occupied by the species in year X |
| SPECIES\_ID | VARCHAR(45) | Species id within a simulation |
| YEAR | INT(11) | [100, 1200000] by =100 |
| sp\_id | INT(11) | Seed id |
| niche\_breadth | VARCHAR(45) | large/narrow |
| dispersal\_ability | VARCHAR(45) | good/poor |

## Niche\_breadth

This table is used to store niche breadth (FN and RN) of every species appeared in the simulations.

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Comment |
| id | INT(11) |  |
| FN\_MIN\_TEMP\_MIN | INT(11) |  |
| FN\_MIN\_TEMP\_MAX | INT(11) |  |
| FN\_MAX\_TEMP\_MIN | INT(11) |  |
| FN\_MAX\_TEMP\_MAX | INT(11) |  |
| FN\_MAX\_PREC\_MIN | INT(11) |  |
| FN\_MAX\_PREC\_MAX | INT(11) |  |
| RN\_MIN\_TEMP\_MIN | INT(11) |  |
| RN\_MIN\_TEMP\_MAX | INT(11) |  |
| RN\_MAX\_TEMP\_MIN | INT(11) |  |
| RN\_MAX\_TEMP\_MAX | INT(11) |  |
| RN\_MAX\_PREC\_MIN | INT(11) |  |
| RN\_MAX\_PREC\_MAX | INT(11) |  |
| SPECIES\_ID | VARCHAR(255) |  |
| YEAR | INT(11) |  |
| sp\_id | INT(11) |  |
| niche\_breadth | VARCHAR(255) |  |
| dispersal\_ability | VARCHAR(255) |  |
| speciation\_year | INT(11) |  |
| extinction\_threshold | INT(11) |  |
| RN\_V | FLOAT |  |
| FN\_V | FLOAT |  |

## Origin

## Speciation\_Convex

## Speciation\_Convex\_E

## Speciation\_Extinction\_BIOME

## Speciation\_Extinction\_Continent

## Speciation\_Extinction\_Continent2

## Speciation\_Extinction\_Lat

## Speciation\_Extinction\_TT

## Speciation\_extinction\_in\_combined\_environments\_XXX

This table is used to store the number of speciation and extinction events per environmental combination (10 degree min temp by 10 degree max temp by 10 mm prec). XXX might be ‘Convex’, ‘Origin’ or ‘Town’.

|  |  |  |
| --- | --- | --- |
| **Name** | **Data Type** | **Comment** |
| ID | Numeric | identity |
| Band\_T\_MIN | Numeric | band of minimal temperature |
| Band\_T\_MAX | Numeric | band of maximal temperature |
| Band\_P\_MAX | Numeric | band of maximal precipitation |
| Value | Numeric | number of events |
| SE | Varchar | S or E. S=speciation E=Extinction |
| sp\_id | Numeric | Seed id |
| niche\_breadth | Varchar | large/narrow |
| dispersal\_ability | Varchar | good/poor |
| speciation\_year | Numeric | 10000 only |
| extinction\_threshold | Numeric | 0/5 |

## Simulations

## Temperate\_Tropic

This table is used to store the distribution (number of cells) of each species (including the original and new species) within temperate or tropic. (18 degree annual temperature). 0 is temperate. 1 is tropic.

|  |  |  |
| --- | --- | --- |
| **Name** | **Data Type** | **Comment** |
| ID | Numeric | identity |
| Temperate\_Tropic | Numeric | [0, 1] |
| COUNT\_Temperate\_Tropic | Numeric | Number of the cells occupied by the species in year X |
| SPECIES\_ID | Varchar | Species id within a simulation |
| YEAR | Numeric | [100, 1200000] by =100 |
| sp\_id | Numeric | Seed id |
| niche\_breadth | Varchar | large/narrow |
| dispersal\_ability | Varchar | good/poor |
| speciation\_year | Numeric | 10000 only |
| extinction\_threshold | Numeric | 0/5 |

## Tree