Explore the data with manydata: : CHEAT SHEET

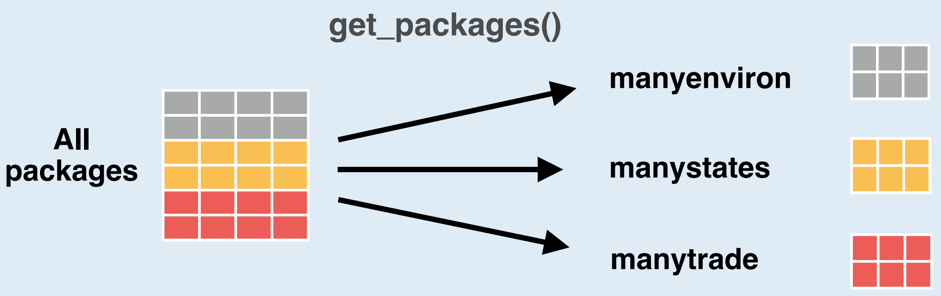
manydata provides the portal to access the packages in the many universe of packages and the data contained in those. **manydata** also contains functions to help users to collect, correct and connect data from different datasets and databases across issue-domains of global governance.

1) Collecting data

get_packages()

```
get_packages(manypackage)
```

A quick and easy way to access and download data from the many packages universe



data_source()

```
data_source(manypackage,
            database,
            dataset)
```

Database from many packages universe

Dataset	Reference
Dataset_A	"Name Surname of authors, year, paper title using the data, publisher, place"
Dataset_B	"Name Surname of authors, year, paper title using the data, publisher, place"
Dataset_C	"Name Surname of authors, year, paper title using the data, publisher, place"

data_contrast()

```
data_contrast(manypackage,
              database,
              dataset)
```

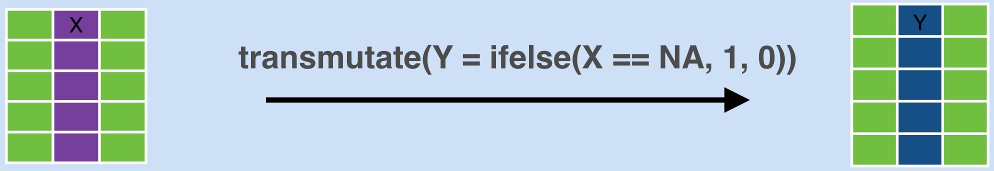
Database

Dataset	Unique ID	Missing data	Rows	Col-umns	Beg	End	URL
Dataset_A	0	3%	3666	8	1351-08-01	NA	https://sourceidatasetA.com
Dataset_B	2765	13%	2765	10	1351-08-01	2020-09-12	https://sourcedatasetB.com
Dataset_C	2390	19%	2390	9	1868-10-17	9999-12-31	https:://sourcedatasetC.com

2) Correcting data

transmutate() keeps unmated variables and drops mutated variables

```
transmutate(data, variable)
```



coalesce_rows() returns first non-missing value for vectors of observations instead of variables

```
coalesce_rows(vector)
```

c(NA, NA, NA, XX, NA) → XX

get_bilaterals() and get_multilaterals() extracts bilateral or multilateral treaties from a full set of treaties in a tibble

```
get_bilaterals(memberships_dataset)
```

CountryID	manyID	Title	Beg	End
H		B		
H		M		
J		B		
C		M		
D		M		
E		M		

get_bilaterals() →

CountryID	manyID	Title	Beg	End
H		B		
J		B		

3) Connecting data

consolidate() allows users to produce a dataset from different datasets within the database. It resolves conflicts in data using the 'resolve' argument, which allows users to specify the value they would like to retain in the dataset.

Use **favour()** to specify the reference dataset for the first NA value before consolidating.

```
consolidate(database,
            rows,
            cols,
            resolve,
            key)
```

Database from a package in the many packages universe

For rows and cols, select either any (all units are retained), or every (only those observations appearing in all parent datasets)

Specify value using 'resolve'

```
resolve = min
```

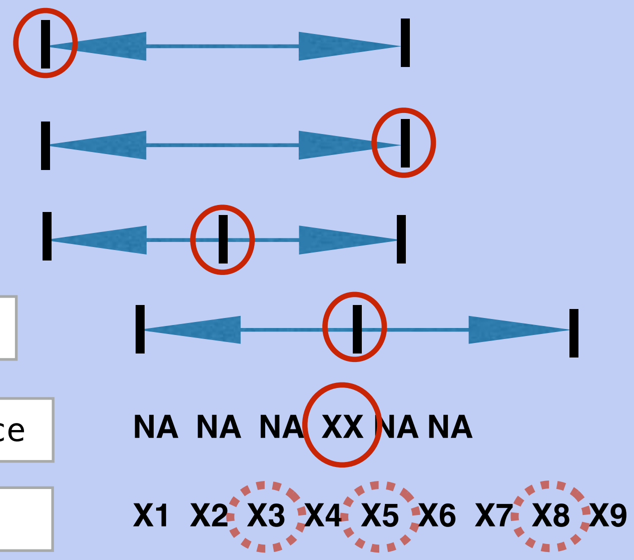
```
resolve = max
```

```
resolve = mean
```

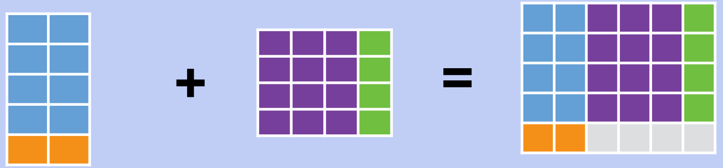
```
resolve = median
```

```
resolve = coalesce
```

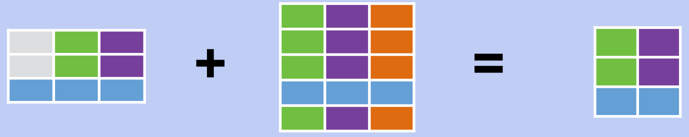
```
resolve = random
```



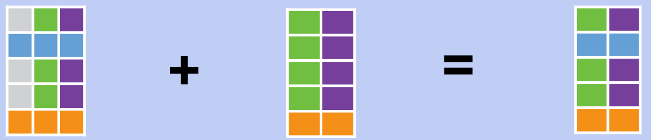
if rows and cols = any



if rows and cols = every



if rows = any & cols = every



CC BY • **http://panarchic.ch** • Learn more at **https://globalgov.github.io/manydata/articles/user.html** • package version 0.7.0 • Updated: 2022-03