

A

Metadata

id	machine_name	date	...	condition	sex	...	p
xxx...xx x	machine_001	2016-09-01	...	A	M	...	p ₁
xxx...xx y	machine_001	2016-09-01	...	B	M	...	p ₂
xxx...xx z	machine_002	2016-09-03	...	A	F	...	p ₃
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	machine _n	date _n	...	condition _n	sex _n	...	p _n

Platform fields

Experiment fields

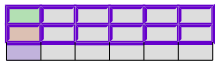
Data

id	t	activity	...	q
xxx...xx x	1	1	⋮	q _{1,1}
xxx...xx x	2	0	⋮	q _{1,2}
xxx...xx x	3	0	⋮	q _{1,3}
xxx...xx x	⋮	⋮	⋮	⋮
xxx...xx y	⋮	⋮	⋮	⋮
xxx...xx z	1	0	⋮	q _{3,1}
xxx...xx z	2	2	⋮	q _{3,2}
xxx...xx z	3	0	⋮	q _{3,3}
xxx...xx z	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
n	⋮	⋮	⋮	q _{n,k_n}

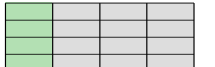
+

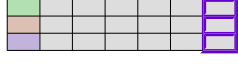
B

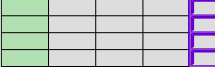
Metadata

Select	dt[CRITERIA, meta = TRUE]  <pre>> male_meta <- dt[sex == "M", meta = TRUE]</pre>
--------	--

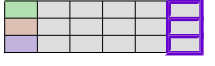
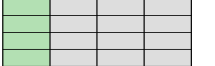

Data

dt[CRITERIA]  <pre>> late_dt <- dt[t > 5]</pre> <p>Note: metadata is updated when selection removes all data from one id.</p>
--

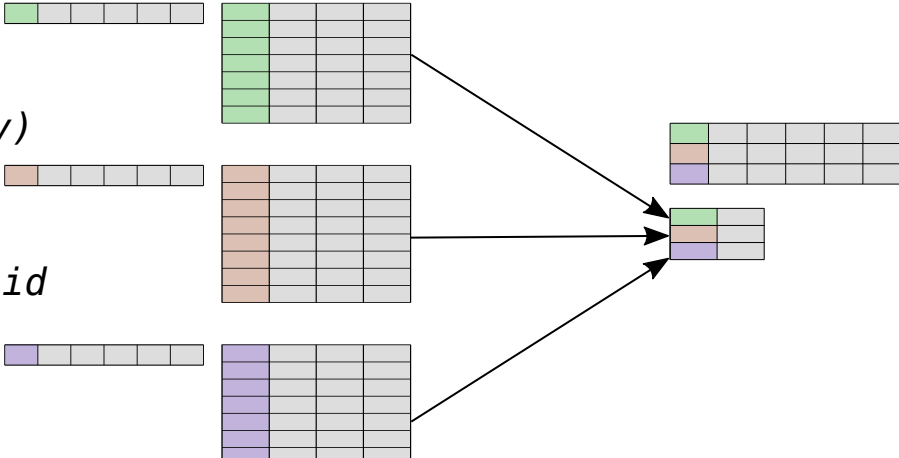
Alter, create & delete (meta)variables	dt[, X := value, meta = TRUE]  <pre>> dt[, genotype := "wt", meta = TRUE] > dt[, sex := NULL] #delete sex</pre>
---	---

dt[, Y := value]  <pre>> dt[, t_2 := t-1] > dt[, t := NULL] #delete t</pre> <p>Note: update data in place. No copy of dt in memory.</p>

Expand metavariables as variables	dt[xmv(X)] <pre>> dt <- dt[xmv(sex) == "M"] > dt[, s := xmv(sex)]</pre>
---	---

Aggregate & summary	dt[, OPERATION, by = id] <pre>> dt <- dt[,.(mean_act = mean(activity)), by = id] > dt[, .N, by = id] # count reads per id</pre>
---------------------	--

OPERATION 
--

Join data & metadata	rejoin(dt) <pre>> full_table <- rejoin(dt)</pre> <p>Note: used mostly after aggregate or preprocessing</p>
----------------------	---

REJOIN 