

~~sd~~
 max 4
 sum 10
 mean 2.5
 [1, 2, 3, 4]

WAT!?

Vectorization
 mutate dep_delay = dep_delay + 1

grouping

dep_delay == max(dep_delay)

dep_delay != 0

month	day	carrier	dep_delay
1	7	EV ①	10
1	10	DL ②	-2
2	20	AA ③	-5
2	23	9E ④	65
3	16	DL	13
3	2	AA	-4
4	26	DL	275
4	5	DL	-4
5	1	AA	-5
5	18	AA	-7
6	23	EV	-3
6	22	EV	-5
7	14	DL	-3
7	10	EV	-6
8	19	DL	4
8	13	DL	-1
9	2	DL	103
9	20	EV	-3
10	26	DL	-4
10	26	AA	1
11	13	EV	-3
11	6	DL	-4
12	23	AA	-2
12	17	AA	-4

10
 4
 4
 30
 30

group_by(month)

How many groups does it create? → 12

group_by(month, carrier)

4 * 12 → 48

group_by(month, carrier)

summarize(meanDelay = mean(dep_delay))

12 AA -3
 11 EV -3
 11 DL -4
 11 AA -1
 10 DL -4

group + filter.
group-by (month) %>%
filter (dep_delay ==
max(dep_delay)
na.rm=T)

group-by (month) %>%
filter (dep_delay != 0)

group-by (month) %>%
mutate (meanDelay = mean(dep_delay)
na.rm=T)

How many rows would this have?

12x