Merging Data

Two datasets

X1	X2
A	1
B	2
C	3

Y1	Y2
A	q
B	r
D	s

Same variable

Note 1: not same order
Note 2: not same values

Many possible outputs

X1	X2
A	1
B	2
C	3

+

Y1	Y2
A	q
B	r
D	s

```
left_join(tbl1, tbl2, by = "X1")
(right_join() not shown)
```

X1	X2	Y2
A	1	q
B	2	r
C	3	NA

```
inner_join(tbl1, tbl2, by = "X1")
```

X1	X2	Y2
A	1	q
B	2	r

Possible outputs

→

```
full_join(tbl1, tbl2, by = "X1")
```

X1	X2	Y2
A	1	q
B	2	r
C	3	NA
D	NA	s

(Using the dplyr package. There is also merge in base R)

Notes

The value we join on (by="") does not need to have the same name in the two tables.

The value we join on can be specified using a combination of multiple columns (album name and artist).

The value we join on does not have to uniquely identify a row, but then we get duplications.

"left join", "inner join", "full outer join" are SQL terminology.