

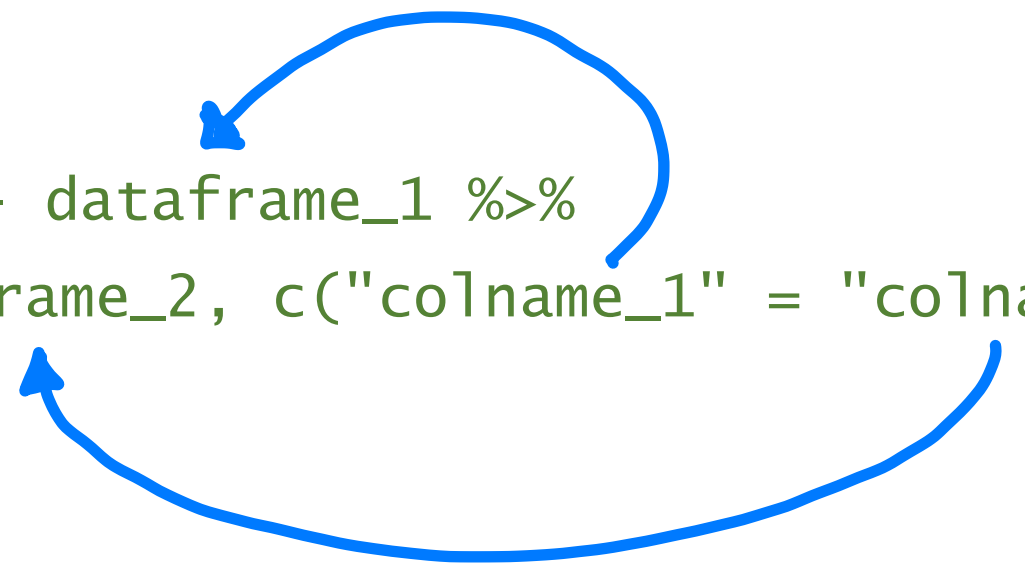
# Joining Datasets

## Part 3

# full\_join() Function

## Command Illustration

```
new_dataframe_name <- dataframe_1 %>%  
  full_join(dataframe_2, c("colname_1" = "colname_2"))
```



The diagram consists of two blue curved arrows. The first arrow starts at the end of the first line of code, `dataframe_1 %>%`, and points to the `full_join` function name on the second line. The second arrow starts at the `dataframe_2` argument on the second line and points back to the `full_join` function name.

## Example 1

Illustration\_Data\_1

Name	Age	num_kids
Val	18	1
Derek	25	0
Whitney	30	2
Daniella	45	1

Illustration\_Data\_2

First_Name	Last_Name	Gender
Val	Chmerkovskiy	Male
Derek	Hough	Male
Whitney	Carson	Female
Sasha	Farber	Male
Daniella	Karagach	Female
Lindsay	Arnold	Female
Mark	Ballas	Male

Full join by these two dataframes by these two columns

**Example 1:** Do a full join between Illustration\_Data\_1 & Illustration\_Data\_2.

```
example_1 <- Illustration_Data_1 %>%  
  full_join(Illustration_Data_2, by = c("Name" = "First_Name"))
```

**Illustration\_Data\_1**

Name	Age	num_kids
Val	18	1
Derek	25	0
Whitney	30	2
Daniella	45	1

**Illustration\_Data\_2**

First_Name	Last_Name	Gender
Val	Chmerkovskiy	Male
Derek	Hough	Male
Whitney	Carson	Female
Sasha	Farber	Male
Daniella	Karagach	Female
Lindsay	Arnold	Female
Mark	Ballas	Male

Output will be a dataframe that looks like:

Name	Age	num_Kids	Last_Name	Gender
Val	18	1	Chmerkovskiy	Male
Derek	25	0	Hough	Male
Whitney	30	2	Carson	Female
Daniella	45	1	Karagach	Female
Sasha	NA	NA	Farber	Male
Lindsay	NA	NA	Arnold	Female
Mark	NA	NA	Ballas	Male

## Example 2

Illustration\_Data\_1

Name	Age	num_kids
Val	18	1
Derek	25	0
Whitney	30	2
Daniella	45	1

Illustration\_Data\_3

Name	Last_Name	Car
Val	Chmerkovskiy	Mercedes
Val	Chmerkovskiy	Tesla
Val	Chmerkovskiy	Audi
Derek	Hough	Ferrari
Lindsay	Arnold	Tesla
Mark	Ballas	BMW

Full join by these two dataframes by these two columns

**Example 2:** Do a full join between Illustration\_Data\_1 & Illustration\_Data\_3.

```
example_2 <- Illustration_Data_1 %>%  
  full_join(Illustration_Data_3, by = c("First_Name" = "Name"))
```

### Illustration\_Data\_1

Name	Age	num_kids
Val	18	1
Derek	25	0
Whitney	30	2
Daniella	45	1

### Illustration\_Data\_3

Name	Last_Name	Car
Val	Chmerkovskiy	Mercedes
Val	Chmerkovskiy	Tesla
Val	Chmerkovskiy	Audi
Derek	Hough	Ferrari
Lindsay	Arnold	Tesla
Mark	Ballas	BMW

Output will be a dataframe that looks like:

Name	Age	num_Kids	Last_Name	Car
Val	18	1	Chmerkovsky	Mercedes
Val	18	1	Chmerkovsky	Tesla
Val	18	1	Chmerkovsky	Audi
Derek	25	0	Hough	Ferrari
Whitney	30	2	NA	NA
Daniella	45	1	NA	NA
Lindsay	NA	NA	Arnold	Tesla
Mark	NA	NA	Ballas	BMW



Will not test students on this!

## Join by Multiply Columns

### Command Illustration

```
new_dataframe <- name_dataframe_x %>%  
  inner_join(name_dataframe_y, by=c("x1" = "y1", "x2" = "y2"))
```

The diagram consists of four red curved arrows. Two arrows originate from the highlighted join conditions, "x1" = "y1" and "x2" = "y2", and point towards the "Command Illustration" header. The other two arrows originate from the same conditions and point towards the "inner\_join" function name in the code line below.

### Example 3

emp\_df

emp_id	name	superior_emp_id	dept_id	dept_branch_id
1	Smith	-1	10	101
2	Rose	1	20	102
3	Williams	1	10	101
4	Jones	2	10	101
5	Brown	2	40	104
<del>6</del>	<del>Brown</del>	<del>2</del>	<del>50</del>	<del>105</del>

dept\_df

dept_id	dept_branch_id	dept_name
10	101	Finance
20	102	Marketing
<del>30</del>	<del>103</del>	<del>Sales</del>
40	104	IT

Dropped  
because  
we  
are  
inner  
joining

**Example 3:** Merge the following two dataframes (emp\_df & dept\_df) by “dept\_id” & “dept\_branch\_id”.

[illegible]

Output will be a dataframe that looks like:

<b>emp_id</b>	<b>name</b>	<b>superior_emp_id</b>	<b>dept_id</b>	<b>dept_branch_id</b>	<b>dept_name</b>
1	Smith	-1	10	101	Finance
2	Rose	1	20	102	Marketing
3	Williams	1	10	101	Finance
4	Jones	2	10	101	Finance
5	Brown	2	40	104	IT