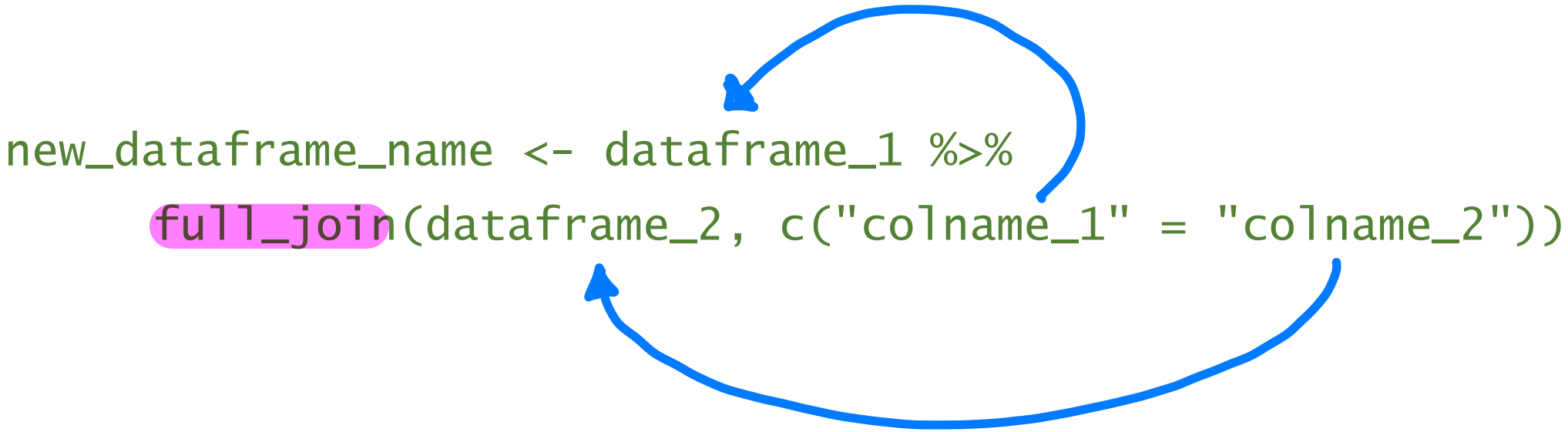
Joining Datasets
Part 3

full_join() Function

Command Illustration



Example 1

Full join by these two dataframes by these two columns

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

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

Full join by these two dataframes by these two columns

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

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:

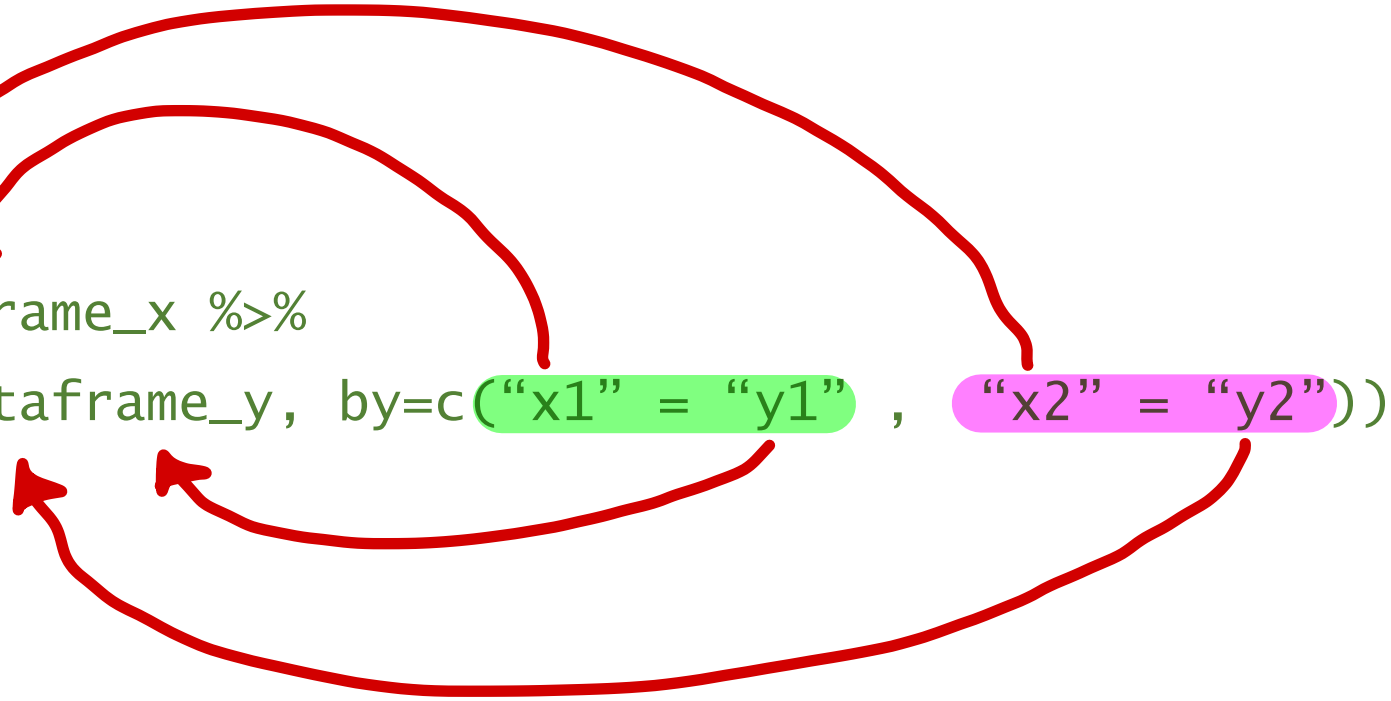
Name	Age	num_Kids	Last_Name	Car
Val	18	1	Chmerkovsky	Mercedes
Val	18	1	Chmerkovsky	Tesla
Val	18	1	Chmerkovsky	Audi
Derek	25	0	Hough	Ferrari
Whitney	30	2	NA	NA
Daniella	45	1	NA	NA
Lindsay	NA	NA	Arnold	Tesla
Mark	NA	NA	Ballas	BMW

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"))
```



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
~~6~~	~~Brown~~	~~2~~	~~50~~	~~105~~

Dropped because we are inner joining

dept_df

dept_id	dept_branch_id	dept_name
10	101	Finance
20	102	Marketing
~~30~~	~~103~~	~~Sales~~
40	104	IT

Example 3: Merge the following two dataframes (emp_df & dept_df) by “dept_id” & “dept_branch_id”.

```
example_3 <- emp_df %>%
   inner_join(dept_df,  by  =  c(“dept_id”  =  “dept_id”,
                      “dept_branch_id” = “dept_branch_id”))
```

Output will be a dataframe that looks like:

emp_id	name	superior_emp_id	dept_id	dept_branch_id	dept_name
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