

Объединение таблиц

Ключи

`head(dataframe)` # первые 6 строк

`tail(dataframe)` # последние 6 строк

`str(dataframe)` # структура данных

`summary(dataframe)` # обзор данных

`dim(dataframe)` # количество строк и столбцов

`names(dataframe)` # названия столбцов

Join Functions

`left_join()`



`right_join()`



`inner_join()`



`full_join()`



Join Functions

`left_join(bands, artists, by = c('left' = 'right'))`

`right_join(artists, bands, by = c('left' = 'right'))`

`inner_join(songs, albums, by = "album")`

`full_join(artists, bands, by = c('left' = 'right'))`

Filtering Join Functions

`semi_join()`



`anti_join()`



Filtering Join Functions

```
semi_join(albums, songs, by = c('left' = 'right'))
```

```
anti_join(artists, bands, by = c('left' = 'right'))
```

Errors

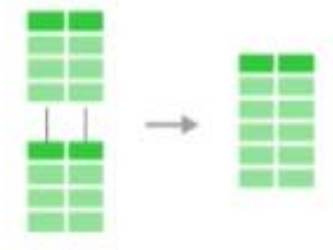
	Missing	Duplicate
key values	Missing key values	Duplicate key columns
key columns	Missing key columns	Duplicate key columns

Binding Functions

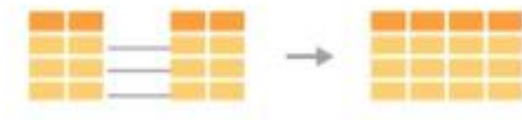
`side_one %>% bind_rows(side_two)`

`bands %>% bind_cols(albums)`

`bind_rows()`



`bind_cols()`



CASE STUDY - Мутации в раке кишечника

Вопрос: Отличается ли количество мутаций определенных генов по полу пациентов и по гистологическому подтипу?

Дано: данные мутаций - 2 таблицы, данные пациентов, данные полученных образцов

Надо: сделать столбчатые диаграммы по количеству мутаций по полу пациентов, разделенные по гистологическим подтипам

CASE STUDY - Мутации в раке кишечника

1. Получить таблицу со следующими клиническими данными - Идентификатор образца, Пол, Гистологический подтип (HISTOLOGICAL_SUBTYPE)
2. Получить полную таблицу со всеми мутациями
3. Найти мутации по гену APC
4. Построить столбчатую диаграмму по полу и по гистологическому подтипу
5. Превратить код в функцию и попробовать на других генах - TP53 и SYNE1