

SQL Analogies in R

Some things you may know how to do in SQL which you can also do in R.

Data:

```
set.seed(2)

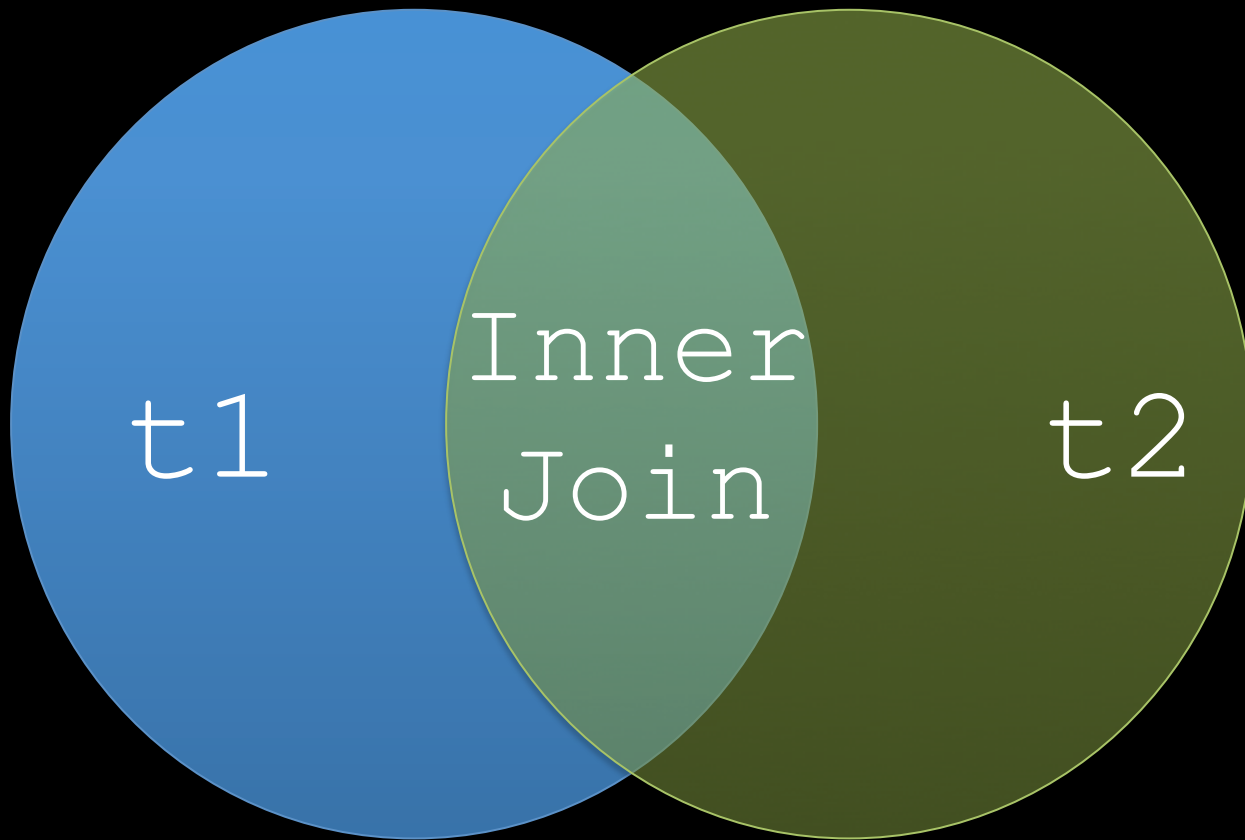
t1 <- data.frame(key1 = sample(letters, 100, replace=TRUE),
                  value1 = rnorm(100))

t2 <- data.frame(key1 = letters,
                  value2 = 1:26)
```

```
> head(t1)
  key1      value1
1    e -0.8382871
2    s  2.0663014
3    o -0.5622471
4    e  1.2757155
5    y -1.0475726
6    y -1.9658782
```

```
> head(t2)
  key1 value2
1    a      1
2    b      2
3    c      3
4    d      4
5    e      5
6    f      6
```

Inner Join



Inner Join

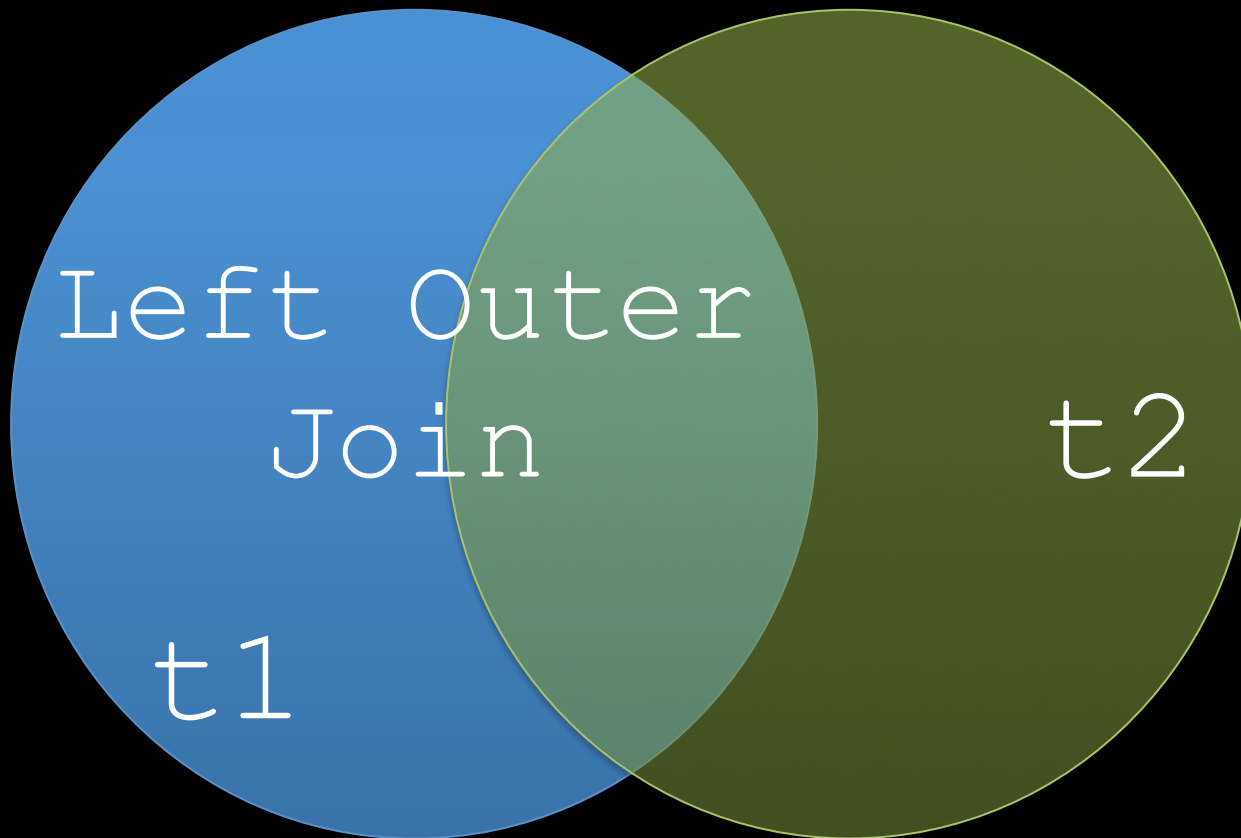
SQL:

```
SELECT t1.*, t2.*  
FROM t1  
INNER JOIN t2 on  
t1.key1 = t2.key1
```

R:

```
merge(t1, t2)
```

Left Outer Join



Left Outer Join

SQL:

```
SELECT t1.*, t2.*  
FROM t1 LEFT OUTER  
JOIN t2  
on t1.key1 = t2.key1
```

R:

```
merge(t1, t2, all.x=TRUE)
```

Select on Criteria

SQL:

```
SELECT *  
FROM t1  
WHERE key1='t'
```

R:

```
subset(t1, key1=='t')
```

OR

```
t1[ which( t1$key1 == "t"), ]
```

See also: ? " ["

Use SQL Syntax in R!

```
require(sqldf)

set.seed(2)
t1 <- data.frame(key1 = sample(letters, 100,
replace=TRUE),
                  value1 = rnorm(100))
t2 <- data.frame(key1 = letters,
                  value2 = 1:26)

## inner join
sqldf("SELECT t1.*, t2.*
      FROM t1 INNER JOIN t2 on t1.key1 = t2.key1
      ")
```