
Recipes 데이터 분석

강사 : 문성민

안동권씨 성화보 데이터 다루기

안동권씨 성화보 데이터 다루기

1번째 과제,

안동권씨 성화보 데이터에서 last name, first name, string에 해당하는 데이터만을 500개 추출하여 아래와 같은 데이터 형태로 만드시오.

```
> head(Re_df_500)
  lastname firstname string
1       권         행     1
2       권       인행    11
3       권         책   111
4       권       균한  1111
5       권       자평  11111
6       권       선개 111111

> str(Re_df_500)
'data.frame':   500 obs. of  3 variables:
 $ lastname : Factor w/ 68 levels "x","元朝","無",...: 9 9 9 9 9 9 9 9 9 9 ...
 $ firstname: Factor w/ 2508 levels "", "00", "000",...: 2296 1712 2082 286 1784 885 438 1641 1971 2178 ...
 $ string   : Factor w/ 6007 levels "1","11","111",...: 1 2 3 4 5 6 7 8 9 5964 ...
```

안동권씨 성화보 데이터 다루기

2번째 과제,

1번과제를 통해 생성된 데이터를 사용하여 직계여부에 대한 데이터를 생성하시오.
(설명 : string은 해당 인물의 족보 관계를 나타내며, string에서 마지막 숫자를 제거하면 해당 인물의 아버지에 해당하는 string이 된다. 타겟이되는 인물이 아버지와 같은 성을 가지고 있으면 0을 다른 성을 가지고 있으면 1을 입력하여 'line'이라는 이름의 변수를 생성하시오.)

```
> head(Re_df_final)
```

```
  lastname firstname string line
1     권       행       1     0
2     권     인행     11     0
3     권       책    111     0
4     권     균한   1111     0
5     권     자평  11111     0
6     권     선개 111111     0
```

```
> str(Re_df_final)
```

```
'data.frame':  500 obs. of  4 variables:
 $ lastname : Factor w/ 68 levels "x","元朝","無",...: 9 9 9 9 9 9 9 9 9 9 ...
 $ firstname: Factor w/ 2508 levels "", "00", "000",...: 2296 1712 2082 286 1784 885 438 1641 1971 2178 ...
 $ string   : Factor w/ 6007 levels "1","11","111",...: 1 2 3 4 5 6 7 8 9 5964 ...
 $ line     : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
```