

Table Conversion : Wide Table <=> Long Table

Table Conversion : Wide Table <=> Long Table

1. tidy package installation
2. 예제 1 : table_conversion__01.csv (Long Table)
 - 2-1. Table Conversion : Long => Wide
 - 2-2. Table Conversion : Wide => Long
3. 예제 2 : table_conversion__02.csv (Long Table)
 - 3-1. Table Conversion : Long => Wide
4. 예제 3 : table_conversion__03.csv (Long Table)
 - 4-1. Table Conversion : Long to Wide
 - 4-2. Table Conversion : Wide to Long

1. tidy package installation

```
install.packages("tidyr")
library(tidyr)
```

2. 예제 1 : table_conversion__01.csv (Long Table)

```
data_table <- read.csv(file.choose(), header=T)
head(data_table, 15)
```

##	행정구역.시군구.별	항목	X2019..03.월
## 1	서울특별시	총전입 [명]	135592
## 2	서울특별시	총전출 [명]	137242
## 3	서울특별시	순이동 [명]	-1650
## 4	서울특별시 시도내이동-시군구내 [명]		37770
## 5	서울특별시 시도내이동-시군구간 전입 [명]		51496
## 6	서울특별시 시도내이동-시군구간 전출 [명]		51496
## 7	서울특별시	시도간전입 [명]	46326
## 8	서울특별시	시도간전출 [명]	47976
## 9	부산광역시	총전입 [명]	36425
## 10	부산광역시	총전출 [명]	38345
## 11	부산광역시	순이동 [명]	-1920
## 12	부산광역시 시도내이동-시군구내 [명]		10605
## 13	부산광역시 시도내이동-시군구간 전입 [명]		15203
## 14	부산광역시 시도내이동-시군구간 전출 [명]		15203
## 15	부산광역시	시도간전입 [명]	

2-1. Table Conversion : Long => Wide

key=항목, value=X2019..03.월

단, '시도내이동-시군구내[명]'에 중복값이 있어서는 안됨.

```
wide_table <- spread(data_table, 항목, X2019..03.월)
head(wide_table)
```

```
## 행정구역.시군구.별 순이동[명] 시도간전입[명] 시도간전출[명]
## 1 광주광역시 -397 6388 6785
## 2 대구광역시 -2425 7429 9854
## 3 대전광역시 -1193 7510 8703
## 4 부산광역시 -1920 10617 12537
## 5 서울특별시 -1650 46326 47976
## 6 세종특별자치시 2076 4607 253
```

2-2. Table Conversion : Wide => Long

```
long_table <- gather(wide_table, '시도간전입[명]':'총전출[명]', key="항목",
value="X2019..03.월")
long_table
```

```
## 행정구역.시군구.별 순이동[명] 항목 X2019..03.월
## 1 광주광역시 -397 시도간전입[명] 6388
## 2 대구광역시 -2425 시도간전입[명] 7429
## 3 대전광역시 -1193 시도간전입[명] 7510
## 4 부산광역시 -1920 시도간전입[명] 10617
## 5 서울특별시 -1650 시도간전입[명] 46326
## 6 세종특별자치시 2076 시도간전입[명]
```

3. 예제 2 : table_conversion__02.csv (Long Table)

```
data_table <- read.csv(file.choose(), header=T)
head(data_table)
```

```
## 행정구역.시군구.별 항목 X2019..03.월
## 1 서울특별시 총전입[명] 135592
## 2 서울특별시 총전출[명] 137242
## 3 서울특별시 순이동[명] -1650
## 4 서울특별시 시도내이동-시군구내[명] 37770
## 5 서울특별시 시도내이동-시군구간 전입[명] 51496
## 6 서울특별시 시도내이동-시군구간 전출[명] 51496
```

3-1. Table Conversion : Long => Wide

key=항목, value=X2019..03.월

```
wide_table <- spread(data_table, 항목, X2019..03.월)
```

단, '시도내이동-시군구내[명]'에 중복값이 있어서 에러가 발생됨.

```
## 에러: Each row of output must be identified by a unique combination of keys.
Keys are shared for 16 rows:
## * 11, 27
## * 15, 31
## * 16, 32
## * 13, 29
## * 14, 30
## * 12, 28
## * 9, 25
## * 10, 26
##
## Call `rlang::last_error()` to see a backtrace
```

wide_table

```
## 행정구역.시군구.별 순이동[명] 시도간전입[명] 시도간전출[명]
## 1 광주광역시 -397 6388 6785
## 2 대구광역시 -2425 7429 9854
## 3 대전광역시 -1193 7510 8703
## 4 부산광역시 -1920 10617 12537
## 5 서울특별시 -1650 46326 47976
## 6 세종특별자치시 2076 4607 2531
```

- Data Table 수정

이 경우 data file의 '시도내이동-시군구내[명]'에 있는 중복값을 수정해야 한다.

수정된 파일 : table_conversion_03.csv

4. 예제 3 : table_conversion__03.csv (Long Table)

```
data_table <- read.csv(file.choose(), header=T)
head(data_table)
```

	행정구역.시군구.별	항목	X2019..03.월
1	서울특별시	총전입[명]	135592
2	서울특별시	총전출[명]	137242
3	서울특별시	순이동[명]	-1650
4	서울특별시	시도내이동-시군구내[명]	37770
5	서울특별시	시도내이동-시군구간 전입[명]	51496
6	서울특별시	시도내이동-시군구간 전출[명]	51496

4-1. Table Conversion : Long to Wide

key=항목, value=X2019..03.월

```
wide_table <- spread(data_table, 항목, X2019..03.월)
wide_table
```

```
## 행정구역.시군구.별 순이동[명] 시도간전입[명] 시도간전출[명]
## 1 대구광역시 -2425 7429 9854
## 2 대구광역시 중구 -171 370 364
## 3 서울특별시 -1650 46326 47976
## 4 서울특별시 중구 -57 615 562
```

4-2. Table Conversion : Wide to Long

```
long_table <- gather(wide_table, '시도간전입[명]':'총전출[명]', key="항목",
value="X2019..03.월")
long_table
```

	행정구역.시군구.별 순이동[명]	항목	X2019..03.월
1	대구광역시	-2425 시도간전입[명]	7429
2	대구광역시 중구	-171 시도간전입[명]	370
3	서울특별시	-1650 시도간전입[명]	46326
4	서울특별시 중구	-57 시도간전입[명]	615
5	대구광역시	-2425 시도간전출[명]	9854
6	대구광역시 중구	-171 시도간전출[명]	

