

```
In [4]: library(dplyr)#그룹화에 필요
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats' :

filter, lag

The following objects are masked from 'package:base' :

intersect, setdiff, setequal, union

내국인 방문인구추출

```
In [1]: #system('gdown --id 1-JkTIm_oydTVMdnXztXY1Dfj1vtVux1q') #내국인(colab 실행)
```

```
In [2]: kor<-read.csv('02-1_내국인유동인구.csv', header = TRUE, fileEncoding = "CP949", encod
```

```
In [5]: kor1<-subset(kor, resd=='그 외')#'그 외'만 추출
head(kor1)
```

A data.frame: 6 × 11

	base_date	time	city	emd_cd	emd_nm	sex	age	resd_pop_cnt	work_pop_cnt	visit
	<chr>	<int>	<chr>	<int>	<chr>	<chr>	<int>	<dbl>	<dbl>	
2	2018-01-01	1	제주 시	50110250	한림읍	여성	0	0	3.7249	
4	2018-01-01	1	제주 시	50110250	한림읍	여성	10	0	4.2577	
6	2018-01-01	1	제주 시	50110250	한림읍	여성	20	0	8.9235	
8	2018-01-01	1	제주 시	50110250	한림읍	여성	30	0	0.0000	
10	2018-01-01	1	제주 시	50110250	한림읍	여성	40	0	13.9136	
12	2018-01-01	1	제주 시	50110250	한림읍	여성	50	0	36.0889	

```
In [22]: kor2<-kor1%>%
  group_by(base_date,city,emd_nm)%>%           #날짜, 도시, 행정동명 기준으로 그룹화
  summarize(visit_pop_cnt=mean(visit_pop_cnt)) #하루 치 평균
head(kor2)
```

`summarise()` has grouped output by 'base_date', 'city'. You can override using the `groups` argument.

A grouped_df: 6 × 4

base_date	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2018-01-01	서귀포시	남원읍	248.40392
2018-01-01	서귀포시	대륜동	218.21885
2018-01-01	서귀포시	대정읍	147.62046
2018-01-01	서귀포시	대천동	141.58808
2018-01-01	서귀포시	동홍동	108.06229
2018-01-01	서귀포시	서홍동	44.55432

```
In [23]: y<-cbind(kor2,date=substr(kor2$base_date,1,7)) #년,월 추출
kor3<-y%>%
  group_by(date,city,emd_nm)%>%
  summarize(visit_pop_cnt=sum(visit_pop_cnt)) #월 별 합계
head(kor3)
tail(kor3)
```

`summarise()` has grouped output by 'date', 'city'. You can override using the `.groups` argument.

A grouped_df: 6 × 4

date	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2018-01	서귀포시	남원읍	7650.916
2018-01	서귀포시	대륜동	7964.611
2018-01	서귀포시	대정읍	4994.301
2018-01	서귀포시	대천동	4671.394
2018-01	서귀포시	동홍동	3881.022
2018-01	서귀포시	서홍동	1671.296

A grouped_df: 6 × 4

date	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2021-06	제주시	일도2동	2601.794
2021-06	제주시	조천읍	13478.916
2021-06	제주시	추자면	1043.922
2021-06	제주시	한경면	3886.112
2021-06	제주시	한림읍	11498.612
2021-06	제주시	화북동	2410.348

```
In [24]: kor4<-rename(kor3, y_m=date,city=city,location=emd_nm)
head(kor4)
```

A grouped_df: 6 × 4

y_m	city	location	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2018-01	서귀포시	남원읍	7650.916
2018-01	서귀포시	대륜동	7964.611
2018-01	서귀포시	대정읍	4994.301
2018-01	서귀포시	대천동	4671.394
2018-01	서귀포시	동홍동	3881.022
2018-01	서귀포시	서홍동	1671.296

장기체류 외국인 방문인구 추출

```
In [9]: #system('gdown --id 1-JwIhXzA-iXPPVjQ_vdSM_guGmXTfAuw') #장기외국인 colab용
```

```
In [13]: lf<-read.csv('02-2_장기체류 외국인 유동인구.csv', header = TRUE, fileEncoding = "CP949")
```

```
In [18]: lf1<-subset(lf, resd=='그외')
head(lf1)
```

A data.frame: 6 × 10

	base_date	time	nationality	city	emd_cd	emd_nm	resd_pop_cnt	work_pop_cnt	visit_po
	<chr>	<int>	<chr>	<chr>	<int>	<chr>	<dbl>	<dbl>	<dbl>
3	2018-01-01	1	CHN	제주시	50110250	한림읍	0	7.8695	51
4	2018-01-01	1	DEU	제주시	50110250	한림읍	0	0.0000	8
6	2018-01-01	1	ETC	제주시	50110250	한림읍	0	0.0000	5
9	2018-01-01	1	IDN	제주시	50110250	한림읍	0	71.3755	63
14	2018-01-01	1	LKA	제주시	50110250	한림읍	0	7.8340	0
16	2018-01-01	1	MMR	제주시	50110250	한림읍	0	5.7265	0

```
In [25]: lf2<-lf1%>%
  group_by(base_date, city, emd_nm)%>%
  summarize(visit_pop_cnt=mean(visit_pop_cnt))
head(lf2)
```

`summarise()` has grouped output by 'base_date', 'city'. You can override using the `groups` argument.

A grouped_df: 6 × 4

base_date	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2018-01-01	서귀포시	남원읍	5.803301
2018-01-01	서귀포시	대륜동	6.769393
2018-01-01	서귀포시	대정읍	15.546368
2018-01-01	서귀포시	대천동	6.987409
2018-01-01	서귀포시	동홍동	5.577592
2018-01-01	서귀포시	서홍동	7.726467

In [26]:

```

y<-cbind(lf2,date=substr(lf2$base_date,1,7))
lf3<-y%>%
  group_by(date,city,emd_nm)%>%
  summarize(visit_pop_cnt=sum(visit_pop_cnt))
head(lf3)
tail(lf3)

```

`summarise()` has grouped output by 'date', 'city'. You can override using the `.groups` argument.

A grouped_df: 6 × 4

date	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2018-01	서귀포시	남원읍	469.7003
2018-01	서귀포시	대륜동	371.7235
2018-01	서귀포시	대정읍	410.5112
2018-01	서귀포시	대천동	234.2381
2018-01	서귀포시	동홍동	270.2584
2018-01	서귀포시	서홍동	426.8739

A grouped_df: 6 × 4

date	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2021-06	제주시	일도2동	210.2710
2021-06	제주시	조천읍	377.1124
2021-06	제주시	추자면	424.4692
2021-06	제주시	한경면	603.8101
2021-06	제주시	한림읍	607.9278
2021-06	제주시	화북동	157.5394

In [35]:

```

lf4<-rename(lf3, y_m2=date,city2=city,location2=emd_nm,visit_pop_cnt_lf=visit_pop_cnt)
head(lf4)

```

A grouped_df: 6 × 4

y_m2	city2	location2	visit_pop_cnt_lf
------	-------	-----------	------------------

<chr>	<chr>	<chr>	<dbl>
2018-01	서귀포시	남원읍	469.7003
2018-01	서귀포시	대륜동	371.7235
2018-01	서귀포시	대정읍	410.5112
2018-01	서귀포시	대천동	234.2381
2018-01	서귀포시	동홍동	270.2584
2018-01	서귀포시	서홍동	426.8739

단기체류 외국인 방문인구 추출

In [14]:

```
단기<-read.csv('02-3_단기체류 외국인 유동인구.csv', header = TRUE, fileEncoding = "CP949")
head(단기)
```

A data.frame: 6 × 7

	base_date	time	nationality	city	emd_cd	emd_nm	visit_pop_cnt
	<chr>	<int>	<chr>	<chr>	<int>	<chr>	<dbl>
1	2018-01-01	1	CHN	제주시	50110250	한림읍	66.3170
2	2018-01-01	1	ETC	제주시	50110250	한림읍	21.1308
3	2018-01-01	1	HKG	제주시	50110250	한림읍	3.6771
4	2018-01-01	1	JPN	제주시	50110250	한림읍	5.5962
5	2018-01-01	1	USA	제주시	50110250	한림읍	3.9078
6	2018-01-01	1	CHN	제주시	50110253	애월읍	123.9840

In [29]:

```
단기1<-단기%>%
  group_by(base_date,city,emd_nm)%>%
  summarize(visit_pop_cnt=mean(visit_pop_cnt))
head(단기1)
```

`summarise()` has grouped output by 'base_date', 'city'. You can override using the `.` groups` argument.

A grouped_df: 6 × 4

base_date	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2018-01-01	서귀포시	남원읍	20.77384
2018-01-01	서귀포시	대륜동	10.52381
2018-01-01	서귀포시	대정읍	13.71054
2018-01-01	서귀포시	대천동	11.17195
2018-01-01	서귀포시	동홍동	17.42905
2018-01-01	서귀포시	서홍동	24.02835

```
In [30]: y<-cbind(단기1,y_m=substr(단기1$base_date,1,7))
단기2<-y%>%
  group_by(y_m,city,emd_nm)%>%
  summarize(visit_pop_cnt=sum(visit_pop_cnt))
head(단기2)
tail(단기2)
```

`summarise()` has grouped output by 'y_m', 'city'. You can override using the `.groups` argument.

A grouped_df: 6 × 4

y_m	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2018-01	서귀포시	남원읍	1686.0106
2018-01	서귀포시	대륜동	420.8336
2018-01	서귀포시	대정읍	356.2630
2018-01	서귀포시	대천동	225.5745
2018-01	서귀포시	동홍동	353.6994
2018-01	서귀포시	서홍동	1355.5257

A grouped_df: 6 × 4

y_m	city	emd_nm	visit_pop_cnt
<chr>	<chr>	<chr>	<dbl>
2021-06	제주시	일도2동	11.87709
2021-06	제주시	조천읍	30.73879
2021-06	제주시	추자면	41.81064
2021-06	제주시	한경면	13.83806
2021-06	제주시	한림읍	28.29782
2021-06	제주시	화북동	10.45366

```
In [34]: 단기3<-rename(단기2,y_m3=y_m,city3=city,location3=emd_nm,visit_pop_cnt_sf=visit_pop_c
head(단기3)
```

A grouped_df: 6 × 4

y_m3	city3	location3	visit_pop_cnt_sf
<chr>	<chr>	<chr>	<dbl>
2018-01	서귀포시	남원읍	1686.0106
2018-01	서귀포시	대륜동	420.8336
2018-01	서귀포시	대정읍	356.2630
2018-01	서귀포시	대천동	225.5745
2018-01	서귀포시	동홍동	353.6994
2018-01	서귀포시	서홍동	1355.5257

방문인구 병합

In [36]:

```
병합<-cbind(kor4,l4,단기3)
head(병합)
```

A grouped_df: 6 × 12

y_m	city	location	visit_pop_cnt	y_m2	city2	location2	visit_pop_cnt_lf	y_m3	city3	loc
<chr>	<chr>	<chr>	<dbl>	<chr>	<chr>	<chr>	<dbl>	<chr>	<chr>	
2018-01	서귀포시	남원읍	7650.916	2018-01	서귀포시	남원읍	469.7003	2018-01	서귀포시	
2018-01	서귀포시	대륜동	7964.611	2018-01	서귀포시	대륜동	371.7235	2018-01	서귀포시	
2018-01	서귀포시	대정읍	4994.301	2018-01	서귀포시	대정읍	410.5112	2018-01	서귀포시	
2018-01	서귀포시	대천동	4671.394	2018-01	서귀포시	대천동	234.2381	2018-01	서귀포시	
2018-01	서귀포시	동홍동	3881.022	2018-01	서귀포시	동홍동	270.2584	2018-01	서귀포시	
2018-01	서귀포시	서홍동	1671.296	2018-01	서귀포시	서홍동	426.8739	2018-01	서귀포시	

In [39]:

```
병합1<-subset(병합,select=-c(y_m2,city2,location2,y_m3,city3,location3))
head(병합1)
tail(병합1)
```

A grouped_df: 6 × 6

y_m	city	location	visit_pop_cnt	visit_pop_cnt_lf	visit_pop_cnt_sf
<chr>	<chr>	<chr>	<dbl>	<dbl>	<dbl>
2018-01	서귀포시	남원읍	7650.916	469.7003	1686.0106
2018-01	서귀포시	대륜동	7964.611	371.7235	420.8336
2018-01	서귀포시	대정읍	4994.301	410.5112	356.2630
2018-01	서귀포시	대천동	4671.394	234.2381	225.5745
2018-01	서귀포시	동홍동	3881.022	270.2584	353.6994
2018-01	서귀포시	서홍동	1671.296	426.8739	1355.5257

A grouped_df: 6 × 6

y_m	city	location	visit_pop_cnt	visit_pop_cnt_lf	visit_pop_cnt_sf
<chr>	<chr>	<chr>	<dbl>	<dbl>	<dbl>
2021-06	제주시	일도2동	2601.794	210.2710	11.87709
2021-06	제주시	조천읍	13478.916	377.1124	30.73879
2021-06	제주시	추자면	1043.922	424.4692	41.81064
2021-06	제주시	한경면	3886.112	603.8101	13.83806
2021-06	제주시	한림읍	11498.612	607.9278	28.29782
2021-06	제주시	화북동	2410.348	157.5394	10.45366

In [40]:

```
write.csv(병합1, '방문인구병합.csv', fileEncoding = "CP949")
```