

혼밥족들이 살고  
싶은 곳은?



# 혼밥족들이 살고 싶은 곳은?

- 1인 가구는 저출산, 고령화와 함께 우리 사회 인구 구조를 특징짓는 중요한 키워드임
- 1인 가구는 2015년 27.2%에서 2017년 28.6%로 더욱 증가함(통계청)
- 1인 세대의 성별, 연령별로 세분화하고 거주지 이전 사유와 함께 이전 지역을 살펴봄으로써 1인 세대 특성 별 선호 지역을 확인해 보고자 함

## 데이터

- 세대원별 세대수 (행정안전부)
- 인구 이동 통계(MDIS통계청)

## 데이터 범위

- 2014년1월~2018년 12월 5년간의 서울시 데이터로 범위

# 데이터 전처리

---

# 데이터 전처리

## 가공 전

예: 2014년 연도별 세대원수별 세대수.csv

	A	B	C	D	E	F	G
1	행정구역	2014년01월_전체세대	2014년01월_1인세대	2014년01월_2인세대	2014년01월_3인세대	2014년01월_4인세대	2014년01월_5인세대
2	전국 (1)	20,466,884	6,878,013	4,168,201	3,797,996	4,118,031	1,149,02
3	서울특별시 (1100000000)	4,182,556	1,509,520	811,779	781,253	822,068	200,08
4	서울특별시 종로구 (1111000000)	73,548	34,171	13,213	11,125	10,899	3,12
5	서울특별시 종로구 청운효자동(1111051500)	5,877	2,076	1,114	1,083	1,158	33
6	서울특별시 종로구 사직동(1111053000)	4,477	1,951	884	759	625	18
7	서울특별시 종로구 삼청동(1111054000)	1,531	...	761	277	206	190
3861	제주특별자치도 서귀포시 대천동(5013060000)		3,017	1,161	628	488	457
3862	제주특별자치도 서귀포시 중문동(5013061000)		3,684	1,410	712	596	576
3863	제주특별자치도 서귀포시 예래동(5013062000)		1,504	636	319	199	169

## 가공 후

행정구역코드	기준일	1인세대	2인세대	3인세대	4인세대	5인이상세대	세대합계	1인세대비율
1111051500	2015년 7월 1 일 수요일	2039	1104	1085	1098	425	5,751	35%
1111053000	2015년 7월 1 일 수요일	1870	863	732	622	244	4,331	43%
1111054000	2015년 7월 1 일 수요일	694	273	189	181	84	1,421	49%
1111055000	2015년 7월 1 일 수요일	1563	959	803	800	319	4,444	35%
1111056000	2015년 7월 1 일 수요일	2163	1551	1470	1544	640	7,368	29%
1111057000	2015년 7월 1 일 수요일	865	627	655	760	230	3,137	28%
1111058000	2015년 7월 1 일 수요일	1130	473	341	315	91	2,350	48%
1111060000	2015년 7월 1 일 수요일	965	448	371	302	133	2,219	43%
1111061500	2015년 7월 1 일 수요일	4258	576	306	267	90	5,497	77%
1111063000	2015년 7월 1 일 수요일	1993	509	356	299	101	3,258	61%
1111064000	2015년 7월 1 일 수요일	2167	781	567	402	186	4,104	55%

# 데이터 전처리

1. 열 중심=> 행 중심으로 변경
2. 합계 값(전국, 서울시 등) 존재=> 합계 값 삭제
3. 서울시 외 지역 존재 => 서울시 외 지역 삭제
4. 행정구역명 복잡도(시도+시군구+읍면동+코드)=> [읍면동 코드]로 단순화
5. 5개의 파일로 존재=> 1개의 파일로 통합 (2014년~ 2018년 세대원수별 세대수)

# 데이터 전처리 - 취합할 파일 list 불러오기

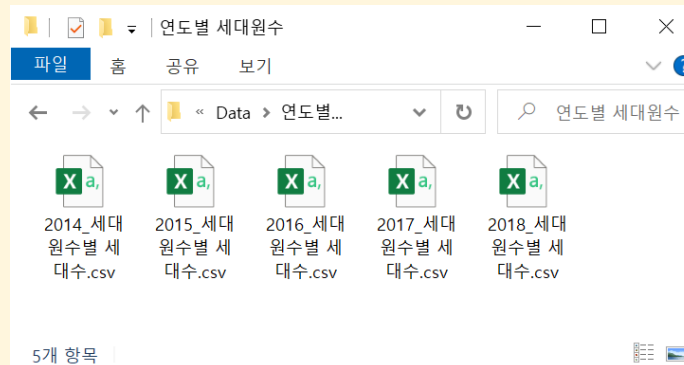
```
: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import re
import os
```

```
: # 취합할 파일 list 불러오기
path='./Data/연도별 세대원수/'
```

```
: files=os.listdir(path)
print(files)
```

['2014\_세대원수별 세대수.csv', '2015\_세대원수별 세대수.csv', '2016\_세대원수별 세대수.csv', '2017\_세대원수별 세대수.csv', '2018\_세대원수별 세대수.csv']

./Data/연도별 세대원수/



# 데이터 전처리 - 파일 읽기

```
for i, file in enumerate(files):
    df=pd.read_csv(path+file, encoding='cp949')

    .
    .
    데이터 전처리
    .
    .
    # concat으로 테이블 합치기
    if i==0:
        res=result_df.copy()
    else:
        res=pd.concat([res,result_df])
```

# 데이터 전처리 - 행정구역 코드 추출

```
df['행정구역코드']=list(map(lambda x: re.findall('[0-9]+',x)[-1], df['행정구역']))
```

	행정구역	2014년01 월_전체세 대	2014년01 월_1인세 대	2014년01 월_2인세 대	2014 월_3 대		2014년 2월_6 !세대	2014 년12 월_7 인세대	2014 년12 월_8 인세대	2014 년12 월_9 인세 대	2014 년12 월_10 인세 대	행정구역코 드
0	전국 (1)	20,466,884	6,878,013	4,168,201	3,79		58,645	62,513	16,286	4,845	4,059	1
1	서울특별시 (1100000000)	4,182,556	1,509,520	811,779	78		42,975	10,198	2,656	846	590	1100000000
2	서울특별시 종로구 (1111000000)	73,548	34,171	13,213	1		732	181	46	21	9	1111000000
3	서울특별시 종로구 청운 효자동 (1111051500)	5,877	2,076	1,114			85	21	2	3	0	1111051500
4	서울특별시 종로구 사직 등 (1111053000)	4,477	1,951	884			51	10	5	2	4	1111053000



# 데이터 전처리 - 집계 행 삭제

# 전국 집계 삭제 - 0번 행 삭제 (전국합계 삭제)

`New_df=df.drop(0)`

# 각 구별 소계 삭제 - '행정구역코드'가 100000배수인 행 삭제

`new_df=new_df[pd.to_numeric(new_df['행정구역코드'])%100000!=0]`

	A	B	C	D	E	F	G	H
1	행정구역	2014년01월 전체세대	2014년01월 1인세대	2014년01월 2인세대	2014년01월 3인세대	2014년01월 4인세대	2014년01월 5인세대	2014년01월 6인세대
2	전국 (1)	20,466,884	6,878,013	4,168,201	3,797,996	4,118,031	1,149,025	266,143
3	서울특별시 (1100000000)	4,182,556	1,509,520	811,779	781,253	822,068	200,084	43,931
4	서울특별시 종로구 (1111000000)	73,548	34,171	13,213	11,125	10,899	3,125	754
5	서울특별시 종로구 동교동(1111051000)	3,077	2,070	1,114	1,003	1,130	334	83
6	서울특별시 종로구 사직동(1111053000)	4,477	1,951	884	759	625	189	50
7	서울특별시 종로구 삼청동(1111054000)	1,531	761	277	206	190	75	14
8	서울특별시 종로구 부암동(1111055000)	1,564	1,607	967	814	867	225	63
9	서울특별시 종로구 평창동(1111056000)	1,921	2,149	1,539	1,444	1,612	480	124
10	서울특별시 종로구 무악동(1111057000)	3,213	880	643	675	765	198	41
11	서울특별시 종로구 교남동(1111058000)	2,645	1,361	496	367	324	76	13
12	서울특별시 종로구 가회동(1111060000)	2,326	1,016	447	380	335	111	28
13	서울특별시 종로구 종로1.2.3.4가동(1111061500)	5,255	3,942	570	359	277	78	20
14	서울특별시 종로구 종로5.6가동(1111063000)	3,108	1,758	517	376	330	103	18
15	서울특별시 종로구 이화동(1111064000)	4,474	2,341	794	624	517	153	32

합계 값 삭제

# 데이터 전처리 - '서울시' 데이터만 추출

#전국데이터 중 "서울시"만 추출

```
s_df=new_df[new_df['행정구역'].str.contains('서울특별시')]
```

행정구역	2014 년01 월_전 체세 대	2014 년01 월_1 인세 대	2014 년01 월_2 인세 대	2014 년01 월_3 인세 대	2014 년01 월_4 인세 대	2014 년01 월_5 인세 대	2014 년01 월_6 인세 대	2014 년01 월_7 인세 대	2014 년01 월_8 인세 대	...	2014 년12 월_2 인세 대	2014 년12 월_3 인세 대	2014 년12 월_4 인세 대	2014 년12 월_5 인세 대	2014 년12 월_6 인세 대	2014 년12 월_7 인세 대	2014 년12 월_8 인세 대	2014 년12 월_9 인세 대	2014 년12 월_10 인세 대	행정구역코 드
3 서울특별시 중 로구 청운효자 동 (1111051500)	5,877	2,076	1,114	1,083	1,158	334	85	23	1	...	1,118	1,093	1,111	313	85	21	2	3	0	1111051500
4 서울특별시 중 로구 사직동 (1111053000)	4,477	1,951	884	759	625	189	50	10	3	...	891	741	636	168	51	10	5	2	4	1111053000
5 서울특별시 중 로구 삼청동 (1111054000)	1,531	761	277	206	190	75	14	7	0	...	273	194	183	67	13	9	0	0	0	1111054000
6 서울특별시 중 로구 부암동 (1111055000)	4,564	1,607	967	814	867	225	63	17	3	...	960	813	830	226	59	19	6	1	0	1111055000

# 데이터 전처리 - 열 중심에서 행중심 테이블로 변환(1)

## 변경 전

행정구역	2014 년01 월_전 체세 대	2014 년01 월_1 인세 대	2014 년01 월_2 인세 대	2014 년01 월_3 인세 대	2014 년01 월_4 인세 대	2014 년01 월_5 인세 대	2014 년01 월_6 인세 대	2014 년01 월_7 인세 대	2014 년01 월_8 인세 대	...	2014 년12 월_2 인세 대	2014 년12 월_3 인세 대	2014 년12 월_4 인세 대	2014 년12 월_5 인세 대	2014 년12 월_6 인세 대	2014 년12 월_7 인세 대	2014 년12 월_8 인세 대	2014 년12 월_9 인세 대	2014 년12 월_10 인세 대	행정구역코 드
3 서울특별시 중 로구 청운효자 동 (1111051500)	5,877	2,076	1,114	1,083	1,158	334	85	23	1	...	1,118	1,093	1,111	313	85	21	2	3	0	1111051500
4 서울특별시 중 로구 사직동 (1111053000)	4,477	1,951	884	759	625	189	50	10	3	...	891	741	636	168	51	10	5	2	4	1111053000
5 서울특별시 중 로구 삼청동 (1111054000)	1,531	761	277	206	190	75	14	7	0	...	273	194	183	67	13	9	0	0	0	1111054000
6 서울특별시 중 로구 부암동 (1111055000)	4,564	1,607	967	814	867	225	63	17	3	...	960	813	830	226	59	19	6	1	0	1111055000

```
s_df=s_df.set_index('행정구역코드').stack().reset_index()  
s_df
```

## 변경 후

	행정구역코드	level_1	0
0	1111051500	2014년01월_전체세대	5,877
1	1111051500	2014년01월_1인세대	2,076
2	1111051500	2014년01월_2인세대	1,114
3	1111051500	2014년01월_3인세대	1,083
4	1111051500	2014년01월_4인세대	1,158
...	...	...	...
55831	1174070000	2014년12월_6인세대	132
55832	1174070000	2014년12월_7인세대	30

# 데이터 전처리 - 열 중심에서 행중심 테이블로 변환(2)

```
# '행정구역' 컬럼 삭제
s_df=s_df.iloc[:,1:]
# 행중심 테이블로 변환
s_df=s_df.set_index('행정구역코드').stack().reset_index()
```

```
s_df.set_index('행정구역코드')
```

행정구역코드	2014년 01월_전체세대	2014년 01월_1인세대	2014년 01월_2인세대	2014년 01월_3인세대	2014년 01월_4인세대	2014년 01월_5인세대	2014년 01월_6인세대
1111051500	5,877	2,076	1,114	1,083	1,158	334	85
1111053000	4,477	1,951	884	759	625	189	50
1111054000	1,531	761	277	206	190	75	14
1111055000	4,564	1,607	967	814	867	225	63
1111056000	7,392	2,149	1,539	1,444	1,612	480	124

```
s_df.set_index('행정구역코드').stack()
```

```
행정구역코드
1111051500 2014년 01월_전체세대      5,877
           2014년 01월_1인세대      2,076
           2014년 01월_2인세대      1,114
           2014년 01월_3인세대      1,083
           2014년 01월_4인세대      1,158
           ...
1174070000 2014년 12월_6인세대       132
           2014년 12월_7인세대        30
           2014년 12월_8인세대         1
           2014년 12월_9인세대         1
           2014년 12월_10인세대        2
Length: 55836, dtype: object
```

```
s_df=s_df.set_index('행정구역코드').stack().reset_index()
s_df
```

	행정구역코드	level_1	0
0	1111051500	2014년 01월_전체세대	5,877
1	1111051500	2014년 01월_1인세대	2,076
2	1111051500	2014년 01월_2인세대	1,114
3	1111051500	2014년 01월_3인세대	1,083
4	1111051500	2014년 01월_4인세대	1,158
...	...	...	...
55831	1174070000	2014년 12월_6인세대	132
55832	1174070000	2014년 12월_7인세대	30

# 데이터 전처리 - 날짜와 세대구분 분할

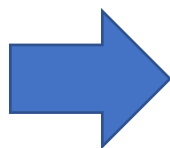
# 기준일 추출

```
s_df['기준일']=list(map(lambda x: x.split('_')[0],s_df['level_1']))
```

# 세대구분 추출

```
s_df['세대']=list(map(lambda x: x.split('_')[1],s_df['level_1']))
```

	행정구역코드	level_1	0
0	1111051500	2014년01월_전체세대	5,877
1	1111051500	2014년01월_1인세대	2,076
2	1111051500	2014년01월_2인세대	1,114
3	1111051500	2014년01월_3인세대	1,083
4	1111051500	2014년01월_4인세대	1,158
...	...	...	...
55831	1174070000	2014년12월_6인세대	132
55832	1174070000	2014년12월_7인세대	30



	행정구역코드	level_1	0	기준일	세대
0	1111051500	2014년01월_전체세대	5,877	2014년01월	전체세대
1	1111051500	2014년01월_1인세대	2,076	2014년01월	1인세대
2	1111051500	2014년01월_2인세대	1,114	2014년01월	2인세대
3	1111051500	2014년01월_3인세대	1,083	2014년01월	3인세대
4	1111051500	2014년01월_4인세대	1,158	2014년01월	4인세대

## 데이터 전처리 - 컬럼 정리

```
# 'level_1' 컬럼 삭제
s_df=s_df.drop('level_1',axis=1)
# 컬럼명 변경(0-> '값')
s_df.columns=['행정구역코드', '값', '기준일', '세대']
# 값 컬럼 데이터형 변환(문자-> 숫자)
s_df['값']=pd.to_numeric(s_df['값'].str.replace(',',''))
```

	행정구역코드	level_1	0	기준일	세대
0	1111051500	2014년01월_전체세대	5,877	2014년01월	전체세대
1	1111051500	2014년01월_1인세대	2,076	2014년01월	1인세대
2	1111051500	2014년01월_2인세대	1,114	2014년01월	2인세대
3	1111051500	2014년01월_3인세대	1,083	2014년01월	3인세대
4	1111051500	2014년01월_4인세대	1,158	2014년01월	4인세대



	행정구역코드	값	기준일	세대
0	1111051500	5877	2014년01월	전체세대
1	1111051500	2076	2014년01월	1인세대
2	1111051500	1114	2014년01월	2인세대
3	1111051500	1083	2014년01월	3인세대
4	1111051500	1158	2014년01월	4인세대

# 데이터 전처리 - 테이블 형 변환(1)

## 변경 전

	행정구역코드	값	기준일	세대
0	1111051500	5877	2014년01월	전체세대
1	1111051500	2076	2014년01월	1인세대
2	1111051500	1114	2014년01월	2인세대
3	1111051500	1083	2014년01월	3인세대
4	1111051500	1158	2014년01월	4인세대

## 변경 후

	행정구역코드	기준일	1인세대	2인세대	3인세대	4인세대	5인이상세대	세대합계	1인세대비율	연도	월
0	1111051500	2014년01월	2076	1114	1083	1158	446	5877	0.35	2014	01
1	1111051500	2014년02월	2084	1113	1079	1168	443	5887	0.35	2014	02
2	1111051500	2014년03월	2083	1116	1088	1174	442	5903	0.35	2014	03
3	1111051500	2014년04월	2084	1114	1078	1169	431	5876	0.35	2014	04
4	1111051500	2014년05월	2086	1113	1079	1164	428	5870	0.36	2014	05

# 데이터 전처리 - 테이블 형 변환(2)

## # 테이블 형 변환

```
result_df=s_df.set_index(['행정구역코드','기준일','세대']).unstack().reset_index()
```

```
s_df.set_index(['행정 구역 코드', '기 준 일', '세 대'])
```

행정구역코드	기준일	값	
		세대	
1111051500	2014년01월	전체세대	5877
		1인세대	2076
		2인세대	1114
		3인세대	1083
		4인세대	1158
...	...	...	...
1174070000	2014년12월	6인세대	132
		7인세대	30

```
s_df.set_index(['행정 구역 코드', '기 준 일', '세 대']).unstack()
```

		값										
	세대	10인세대	1인세대	2인세대	3인세대	4인세대	5인세대	6인세대	7인세대	8인세대	9인세대	전체세대
행정구역코드	기준일											
1111051500	2014년01월	0	2076	1114	1083	1158	334	85	23	1	3	5877
	2014년02월	0	2084	1113	1079	1168	325	90	25	1	2	5887
	2014년03월	0	2083	1116	1088	1174	323	91	24	2	2	5903
	2014년04월	0	2084	1114	1078	1169	315	88	24	2	2	5876

```
result_df=s_df.set_index(['행정 구역 코드', '기 준 일', '세 대']).unstack().reset_index()
```

```
result_df.head()
```

행정구역코드		기준일	값										
세대			10인세대	1인세대	2인세대	3인세대	4인세대	5인세대	6인세대	7인세대	8인세대	9인세대	전체세대
0	1111051500	2014년01월	0	2076	1114	1083	1158	334	85	23	1	3	5877
1	1111051500	2014년02월	0	2084	1113	1079	1168	325	90	25	1	2	5887
2	1111051500	2014년03월	0	2083	1116	1088	1174	323	91	24	2	2	5903
3	1111051500	2014년04월	0	2084	1114	1078	1169	315	88	24	2	2	5876



# 데이터 전처리 - 테이블 형 변환(3)

#컬러명 변경

```
result_df.columns=['행정구역코드', '기준일', '10인세대', '1인세대', '2인세대', '3인세대', '4인세대', '5인세대', '6인세대', '7인세대', '8인세대', '9인세대', '전체세대']
```

변경 전

	행정구역코드	기준일	값										
세대			10인세대	1인세대	2인세대	3인세대	4인세대	5인세대	6인세대	7인세대	8인세대	9인세대	전체세대
0	1111051500	2014년01월	0	2076	1114	1083	1158	334	85	23	1	3	5877
1	1111051500	2014년02월	0	2084	1113	1079	1168	325	90	25	1	2	5887
2	1111051500	2014년03월	0	2083	1116	1088	1174	323	91	24	2	2	5903
3	1111051500	2014년04월	0	2084	1114	1078	1169	315	88	24	2	2	5876
4	1111051500	2014년05월	0	2086	1113	1079	1164	312	90	22	2	2	5870

변경 후

	행정구역코드	기준일	10인세대	1인세대	2인세대	3인세대	4인세대	5인세대	6인세대	7인세대	8인세대	9인세대	전체세대
0	1111051500	2014년01월	0	2076	1114	1083	1158	334	85	23	1	3	5877
1	1111051500	2014년02월	0	2084	1113	1079	1168	325	90	25	1	2	5887
2	1111051500	2014년03월	0	2083	1116	1088	1174	323	91	24	2	2	5903
3	1111051500	2014년04월	0	2084	1114	1078	1169	315	88	24	2	2	5876
4	1111051500	2014년05월	0	2086	1113	1079	1164	312	90	22	2	2	5870

# 데이터 전처리 - 5인 세대 이상 합 구하기

# 5인세대 이상합 구하기

```
result_df['5인이상세대']=result_df['5인세대']+result_df['6인세대']+result_df['7인세대']  
+result_df['8인세대']+result_df['9인세대']+result_df['10인세대']
```

	행정구역코드	기준일	10인세대	1인세대	2인세대	3인세대	4인세대	5인세대	6인세대	7인세대	8인세대	9인세대	전체세대	5인이상세대
0	1111051500	2014년01월	0	2076	1114	1083	1158	334	85	23	1	3	5877	446
1	1111051500	2014년02월	0	2084	1113	1079	1168	325	90	25	1	2	5887	443
2	1111051500	2014년03월	0	2083	1116	1088	1174	323	91	24	2	2	5903	442
3	1111051500	2014년04월	0	2084	1114	1078	1169	315	88	24	2	2	5876	431
4	1111051500	2014년05월	0	2086	1113	1079	1164	312	90	22	2	2	5870	428

## 데이터 전처리 - 컬럼 정리

# 필요한 열만 추출

```
result_df=result_df[['행정구역코드','기준일','1인세대','2인세대','3인세대','4인세대','5인이상세대']]
```

	행정구역코드	기준일	1인세대	2인세대	3인세대	4인세대	5인이상세대
0	1111051500	2014년01월	2076	1114	1083	1158	446
1	1111051500	2014년02월	2084	1113	1079	1168	443
2	1111051500	2014년03월	2083	1116	1088	1174	442
3	1111051500	2014년04월	2084	1114	1078	1169	431
4	1111051500	2014년05월	2086	1113	1079	1164	428

## 데이터 전처리 - '세대합계', '1인세대 비율' 구하기

```
result_df['세대합계']=np.sum(result_df.iloc[:,2:].values,axis=1)
result_df['1인세대비율']=round(result_df['1인세대']/result_df['세대합계'],ndigits=2)
```

	행정구역코드	기준일	1인세대	2인세대	3인세대	4인세대	5인이상세대	세대합계	1인세대비율
0	1111051500	2014년01월	2076	1114	1083	1158	446	5877	0.35
1	1111051500	2014년02월	2084	1113	1079	1168	443	5887	0.35
2	1111051500	2014년03월	2083	1116	1088	1174	442	5903	0.35
3	1111051500	2014년04월	2084	1114	1078	1169	431	5876	0.35
4	1111051500	2014년05월	2086	1113	1079	1164	428	5870	0.36

## 데이터 전처리 - '기준일'에서 '연도' 와 '월' 분리

```
result_df['연도']=result_df['기준일'].str[:4]  
result_df['월']=result_df['기준일'].str[-3:-1]
```

	행정구역코드	기준일	1인세대	2인세대	3인세대	4인세대	5인이상세대	세대합계	1인세대비율	연도	월
0	1111051500	2014년01월	2076	1114	1083	1158	446	5877	0.35	2014	01
1	1111051500	2014년02월	2084	1113	1079	1168	443	5887	0.35	2014	02
2	1111051500	2014년03월	2083	1116	1088	1174	442	5903	0.35	2014	03
3	1111051500	2014년04월	2084	1114	1078	1169	431	5876	0.35	2014	04
4	1111051500	2014년05월	2086	1113	1079	1164	428	5870	0.36	2014	05

# 데이터 전처리 - concat으로 데이터프레임 합치기

```
if i==0:  
    res=result_df.copy()  
else:  
    res=pd.concat([res,result_df])
```

```
: print(res.shape)  
display(res)
```

(25422, 11)

	행정구역코드	기준일	1인세대	2인세대	3인세대	4인세대	5인이상세대	세대합계	1인세대비율	연도	월
0	1111051500	2014년01월	2076	1114	1083	1158	446	5877	0.35	2014	01
1	1111051500	2014년02월	2084	1113	1079	1168	443	5887	0.35	2014	02
2	1111051500	2014년03월	2083	1116	1088	1174	442	5903	0.35	2014	03
3	1111051500	2014년04월	2084	1114	1078	1169	431	5876	0.35	2014	04
4	1111051500	2014년05월	2086	1113	1079	1164	428	5870	0.36	2014	05
...	...	...	...	...	...	...	...	...	...	...	...
5083	1174070000	2018년08월	3350	2348	2366	2418	672	11154	0.30	2018	08
5084	1174070000	2018년09월	3366	2361	2354	2412	669	11162	0.30	2018	09
5085	1174070000	2018년10월	3379	2377	2340	2401	666	11163	0.30	2018	10
5086	1174070000	2018년11월	3392	2378	2339	2382	659	11150	0.30	2018	11
5087	1174070000	2018년12월	3409	2393	2335	2353	657	11147	0.31	2018	12

25422 rows × 11 columns

# 데이터 시각화

---

# 데이터 시각화 - 연도별 세대수 변화

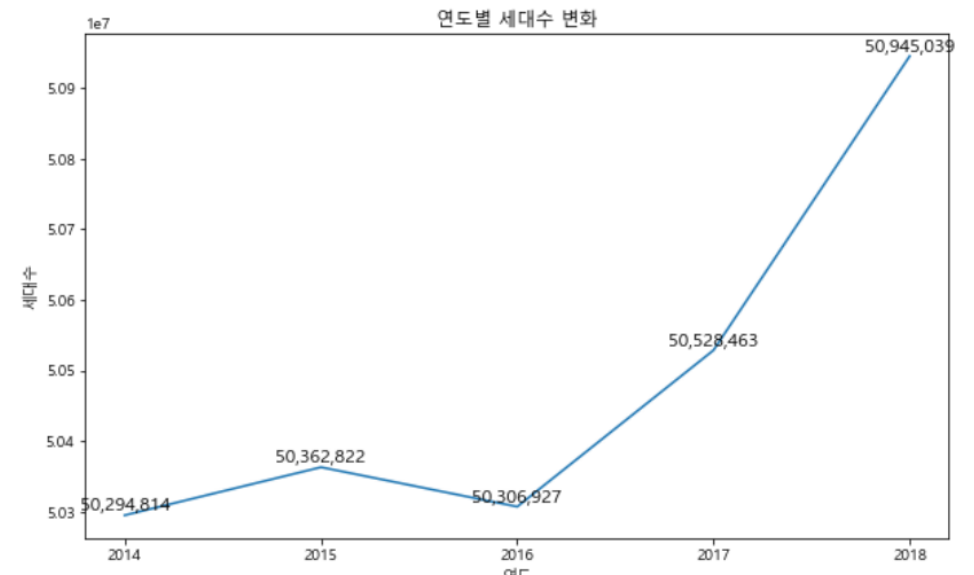
#연도별 세대수 변화

```
plt.rcParams["figure.figsize"] = (10, 6)#차트 전체 기본 크기
```

```
res_year=res.groupby('연도')['세대합계'].sum()
```

```
plt.plot(res_year)
plt.title('연도별 세대수 변화');
plt.xlabel('연도')
plt.ylabel('세대수')
```

```
for i in res_year.index:
    height = res_year[i]
    plt.text(i, height + 100, f'{height:,}', ha='center', va='bottom', size = 12)
```

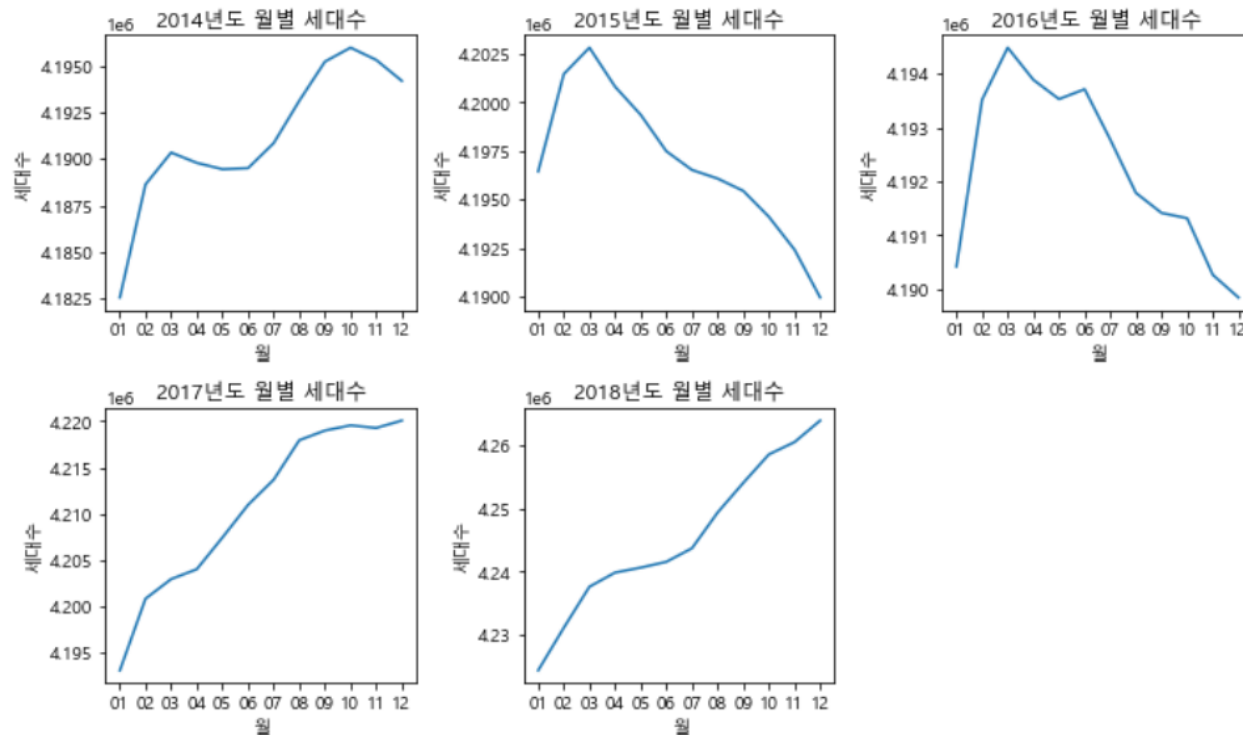




# 데이터 시각화 - 월 년도의 월별 세대수 변화

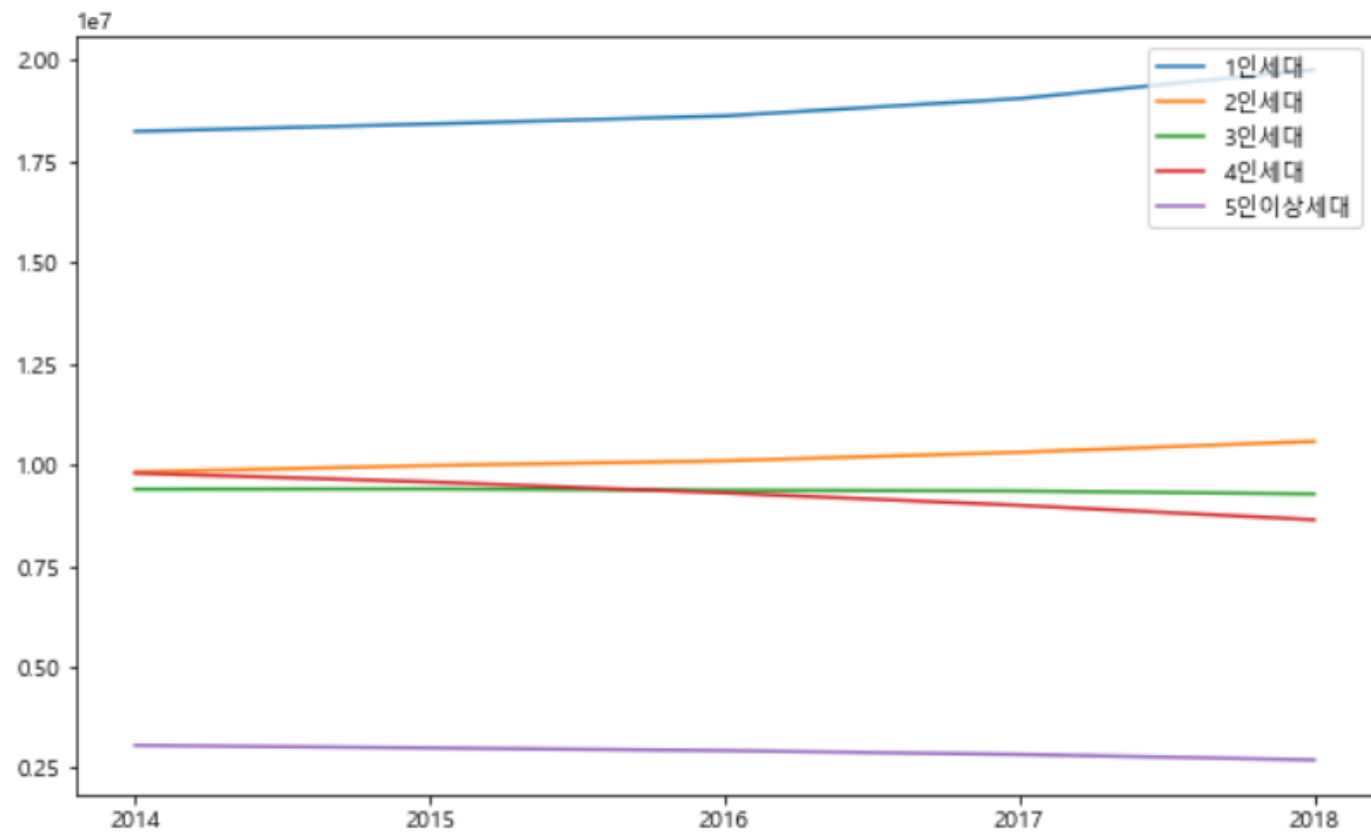
```
for cnt,i in enumerate(res['연도'].unique()):  
    ct=res[res['연도']==i].groupby('월')['세대합계'].sum()
```

```
plt.subplot(2,3,cnt+1)  
plt.plot(ct)  
plt.title(i+'년도 월별 세대수');  
plt.xlabel('월')  
plt.ylabel('세대수')  
plt.tight_layout()
```



# 데이터 시각화 - 세대 형태의 변화 탐색

```
# 세대 구성 형태 변화  
ct=res.groupby('연도').sum().iloc[:, :-2]  
plt.plot(ct)  
plt.legend(ct)
```



# 데이터 시각화 - 세대 형태의 변화 탐색

```
#년도별 세대 구성 형태
```

```
#차트 전체 기본 크기
```

```
plt.rcParams["figure.figsize"] = (15, 8)
```

```
for cnt,i in enumerate(res['연도'].unique()):
```

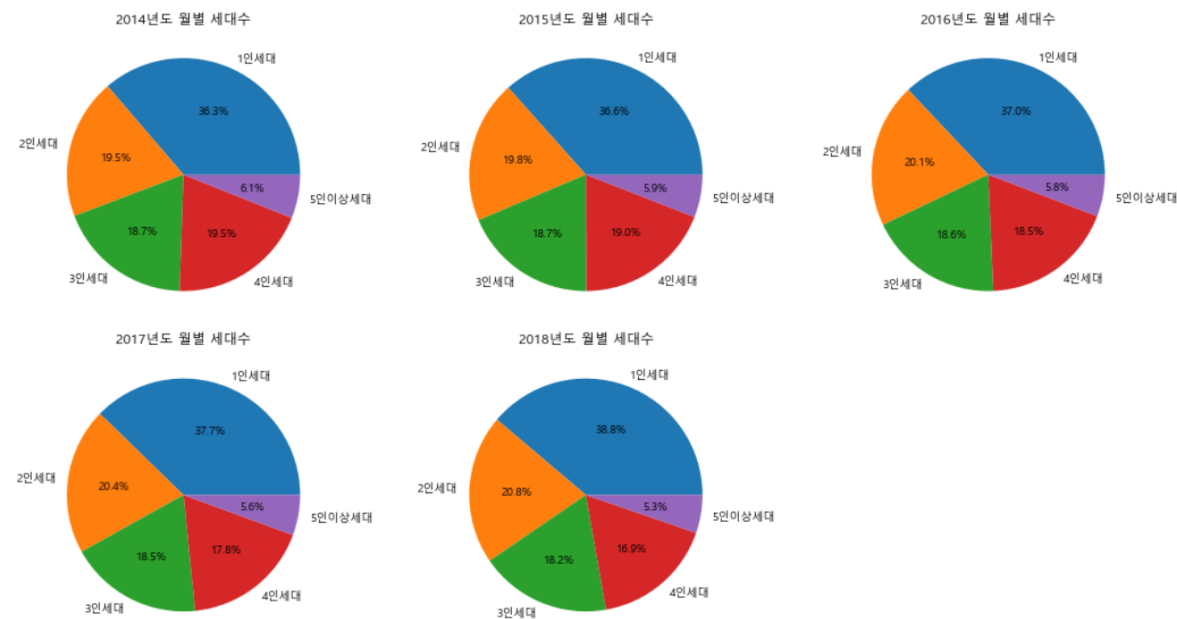
```
    ct=res.groupby('연도').sum().iloc[:,2].T
```

```
    plt.subplot(2,3,cnt+1)
```

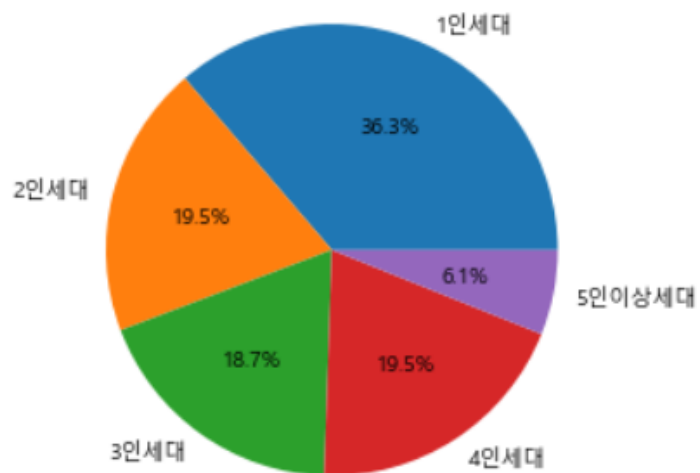
```
    plt.pie(ct[i],labels=ct.index,autopct = '%1.1f%%')
```

```
    plt.title(i+'년도 월별 세대수');
```

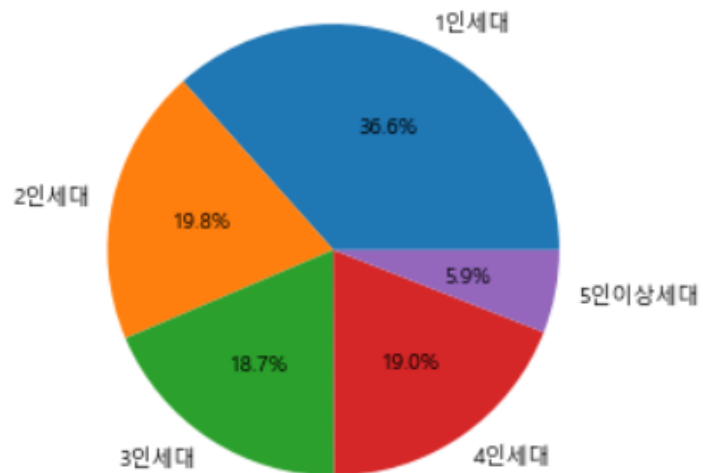
```
plt.tight_layout()
```



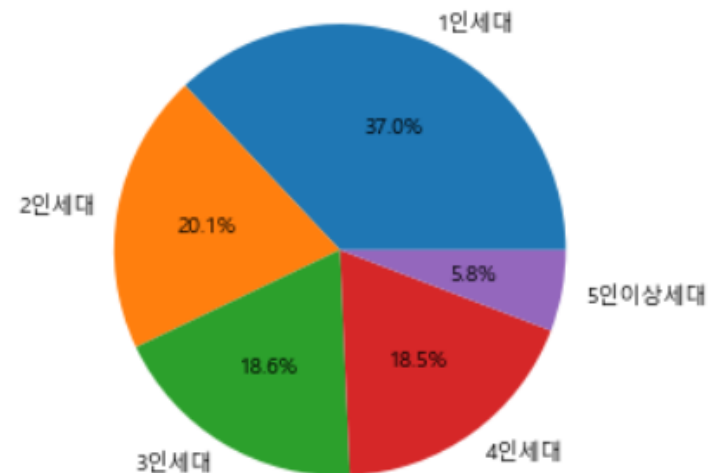
2014년도 월별 세대수



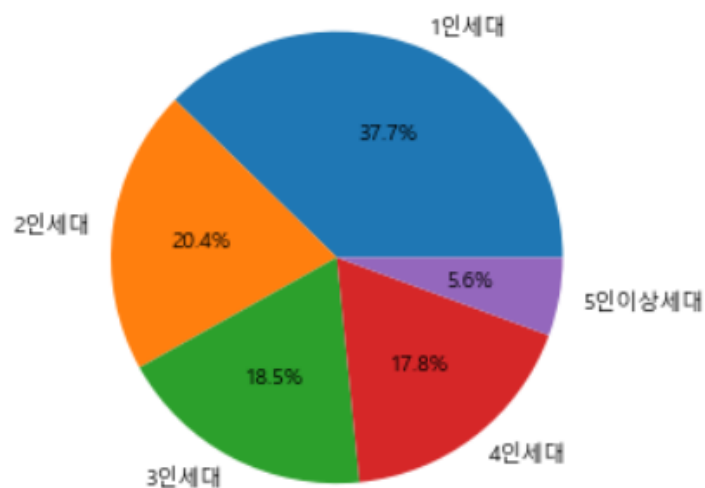
2015년도 월별 세대수



2016년도 월별 세대수



2017년도 월별 세대수



2018년도 월별 세대수

