

R프로그래밍및실습 연습_2

Jinseog Kim
Dongguk University
jskim1986@gmail.com

2018-05-17

- 1 아래의 URL의 데이터는 2015년 우리나라의 시도별/연령대별/성별 인구 데이터이다.
- 2 홈페이지의 데이터 파일 건강보험인구현황.csv/2015_사망사고.csv를 읽어 데이터프레임 pop_2015와 acc_2015을 만들어라.

```
pop_f <- "http://datamining.dongguk.ac.kr/lectures/2018-1/R/건강보험인구현황.csv"
acc_f <- "http://datamining.dongguk.ac.kr/lectures/2018-1/R/2015_사망사고.csv"
cars_f <- "http://datamining.dongguk.ac.kr/lectures/2018-1/R/자동차등록대수(11년-15년).csv"
pop_2015 <- read.csv(pop_f, fileEncoding = "CP949")
acc_2015 <- read.csv(acc_f, fileEncoding = "CP949")
cars_2015 <- read.csv(cars_f, fileEncoding = "CP949")
cars_2015$구분 <- factor(as.character(cars_2015$구분),
                        levels=c("서울", "부산", "대구", "인천", "광주", "대전",
                                "울산", "세종", "경기", "강원", "충북", "충남", "전북", "전남",
                                "경북", "경남", "제주"))
cars_2015 <- subset(cars_2015, select=c("구분", "X2015"))

acc_2015$발생지시도 <- factor(stringr::str_trim(as.character(acc_2015$발생지시도)),
                              levels=c("서울", "부산", "대구", "인천", "광주", "대전",
                                        "울산", "세종", "경기", "강원", "충북", "충남", "전북", "전남",
                                        "경북", "경남", "제주"))

pop_2015$시도 <- as.character(pop_2015$시도)
pop_2015$시도 <- gsub("특별시|광역시|특별자치시|특별자치도|도", "", pop_2015$시도)
do_idx <- nchar(pop_2015$시도)==3
pop_2015$시도[do_idx] <- paste0(substring(pop_2015$시도[do_idx], 1, 1),
                               substring(pop_2015$시도[do_idx], 3, 3))
pop_2015$시도 <- factor(pop_2015$시도, levels=unique(pop_2015$시도))
```

3 시도별 총인구수(단위 100만)를 구하시오

	시도	인구수
1	서울	10.118348
2	부산	3.492557
3	대구	2.496285
4	인천	2.949187
5	광주	1.479490
6	대전	1.522189
7	울산	1.194132
8	세종	0.212984
9	경기	12.733893
10	강원	1.549183
11	충북	1.603385
12	충남	2.111740
13	전북	1.879683
14	전남	1.920319
15	경북	2.726464
16	경남	3.415010
17	제주	0.629575

4 사망사고 중 사고유형_대분류가 '차대차'이고 컬럼명이 '발생지시도', '사망자수'만을 추출하여 acc_2015_차대차에 저장하시오.

	발생지시도	사망자수
1	울산	1
4	경기	1
6	경남	3
9	제주	1
11	경기	1
12	경기	1

5 acc_2015_차대차에서 총 발생건수 및 총 사망자수를 구하시오.

	발생지시도	사고건수	사망자수
1	서울	114	116
2	부산	64	70
3	대구	60	62
4	인천	52	54
5	광주	38	39
6	대전	44	46

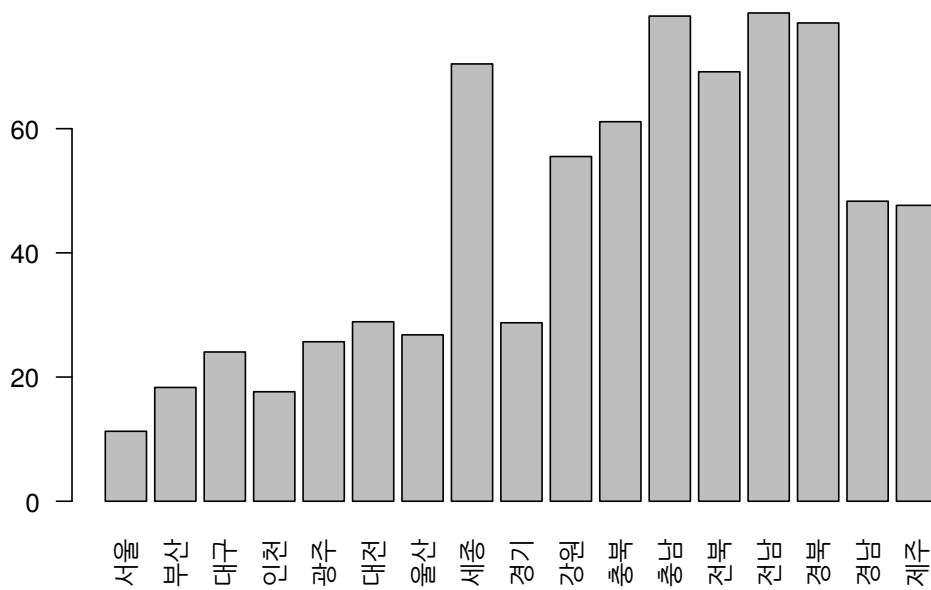
7	울산	32	33
8	세종	15	15
9	경기	366	383
10	강원	86	92
11	충북	98	110
12	충남	165	175
13	전북	130	144
14	전남	151	159
15	경북	210	225
16	경남	165	177
17	제주	30	34

6 시도별 총인구수와 시도별 사망사고 중 차대차로 인한 발생건수 및 사망자수 데이터를 시도별로 병합하라.

	시도	인구수	사고건수	사망자수
1	서울	10.118348	114	116
2	부산	3.492557	64	70
3	대구	2.496285	60	62
4	인천	2.949187	52	54
5	광주	1.479490	38	39
6	대전	1.522189	44	46
7	울산	1.194132	32	33
8	세종	0.212984	15	15
9	경기	12.733893	366	383
10	강원	1.549183	86	92
11	충북	1.603385	98	110
12	충남	2.111740	165	175
13	전북	1.879683	130	144
14	전남	1.920319	151	159
15	경북	2.726464	210	225
16	경남	3.415010	165	177
17	제주	0.629575	30	34

- 7 시도별 총인구수와 시도별 사망사고 중 차대차로 인한 발생건수 및 사망자수 데이터를 시도별로 병합하라.
- 8 시도별 인구대비_사고건수, 인구대비_사망자수, 사고건수대비_사망자수를 구하라.

시도별 인구대비 사고건수



시도별 자동차수대비 사망자수

