

R프로그래밍및실습 마지막 퀴즈_5

2018-06-14

- 인터넷을 비롯한 모든 자료 참조 가능 (단, 이메일, SNS, 토론등 부정행위 적발 시 F처리)
 - 답안지 작성
 - 답안지다운로드: <http://datamining.dongguk.ac.kr/lectures/2018-1/R/답안지.hwp>
 - R 코드, 텍스트 결과, 이미지 캡처
 - 답안지 제출
 - 이메일: jskim1986@gmail.com
 - 제목: “R프로그래밍(마지막 Quiz) - 자신의 이름”
 - 내용: “첨부파일 참조”만 쓸 것
 - 첨부: “답안지.hwp” : 파일명 바꾸지 말 것
 - 각 문항의 배점은 10점
-

1 아래의 R코드를 이용하여 데이터프레임 health_2014_s을 생성하시오.

```
# 홈페이지에서 "건강검진_2014_일부추출.csv" 데이터를 다운로드하여 자신이 지정한 디렉토리에 저장
Data_DIR <- "C:/download/" # 자신이 지정한 디렉토리
health_2014 <- read.csv(paste0(Data_DIR, "건강검진_2014_일부추출.csv"), fileEncoding = "CP949")
student_id <- 1234567 # <=== 여기에 자신의 학번을 입력할 것 (반드시!!!!)
set.seed(student_id)
sidx <- sample(5000, replace = F)
health_2014_s <- health_2014[sidx,]
```

1.1 위의 R코드를 라인별로 설명하시오.

1.2 health_2014_s자료에서 시도별 건강검진 대상자수를 구하라.

1.3 “성별” 변수를 범주형 변수로, 각 수준(level)값이 1은 “남”, 2는 “여”로 수준이름(label)을 바꾸시오.

1.4 흡연상태별 수축기혈압과 이완기혈압의 평균을 막대그래프로 나타내시오.

1.5 시도별로 남자 “40대미만”의 평균 “신장”을 “선그래프”로 나타내시오.

1.6 “성별”, “연령”에 따라 “신장”의 평균 및 표준편차를 구하시오.

1.7 흡연상태가 “1”인 여성의 “수축기혈압”과 “이완기혈압”의 상관계수를 구하시오.

2 아래는 신입생을 대상으로 설문조사한 자료이고 각 변수에 대한 설명은 아래와 같다.

1. 성별: 1.남; 2.여
2. 소속: 단과대학
3. 지원동기: 1. 성적(합격가능성); 2.장학금; 3. 사회적 평판; 4.졸업 후 진로;5.우수한 교육제도 및 시설; 6.학문적 흥미와 적성; 7. 학사교류제도
4. 관심사항: 1. 학업; 2. 취업준비; 3. 동아리활동; 4. 교우관계
5. 학교생활만족도: 1. 매우만족; 2. 만족; 3. 보통; 4. 불만족; 5. 매우불만족

```
# 홈페이지에서 "설문조사.csv" 데이터를 다운로드하여 자신이 지정한 디렉토리에 저장
Data_DIR <- "C:/download/"
survey_2014 <- read.csv(paste0(Data_DIR, "설문조사.csv"),
                        fileEncoding = "CP949")
```

2.1 소속별/성별 조사대상 학생수를 구하라.

2.2 소속별 조사대상 지원동기에 따른 학생수의 비율을 구하라.

2.3 소속별 성별 학교생활만족도의 평균점수와 표준편차를 구하라.