**[프라임칼리지] 통계학개론 - 1차 과제물(100점 만점)**

**제출기한: 2024.04.02.(화) 23:59**

**추가연장제출: 2024.04.03.(수) 23:59**

(‘과제’ 메뉴)

**이 름:\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

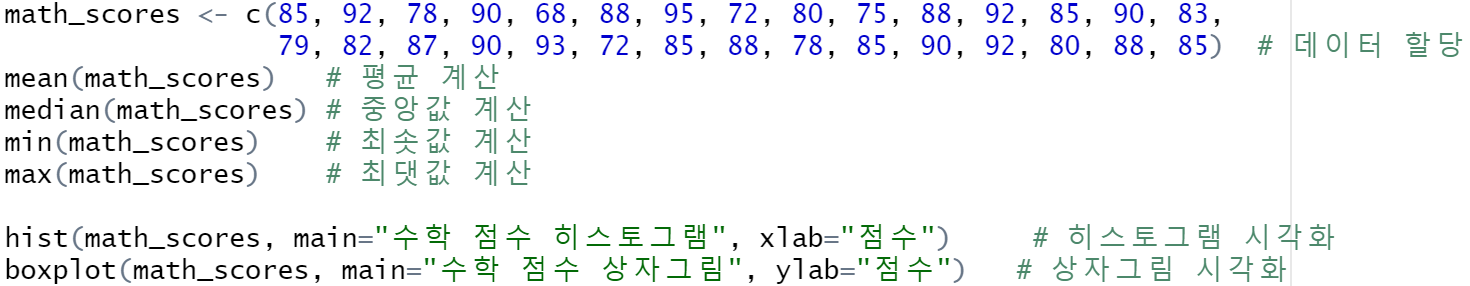
**□ 모든 문항에 대해 답만 쓰지 말고, ‘R 코드’와 ‘R 프로그래밍을 실행한 결과(캡처 이미지)’를 각각 첨부해주세요. (총 4문항)**

- 모든 문항에 대해 결과에 대한 **해석**까지 반드시 포함해주세요.

- **R 실습 과제**이므로 엑셀이나 공학용 계산기, 파이썬 등 다른 프로그램이나 손으로 직접 계산한 답안은 평가가 불가능합니다.

**문제1.** [3강. 데이터 수치요약]

다음은 한 반 학생 30명의 수학 점수 데이터이다. R 프로그램을 사용하여 'math\_scores'라는 변수에 30명의 점수를 할당한 후, 평균, 중앙값, 최솟값, 최댓값을 구하고, 히스토그램과 상자그림으로 나타내시오. 그리고 출력된 기술통계량과 그래프를 통해 알 수 있는 데이터의 특성을 서술하시오. (25점)



\* 참고: 데이터 입력의 편의를 위해 데이터는 아래와 같이 텍스트로 제공

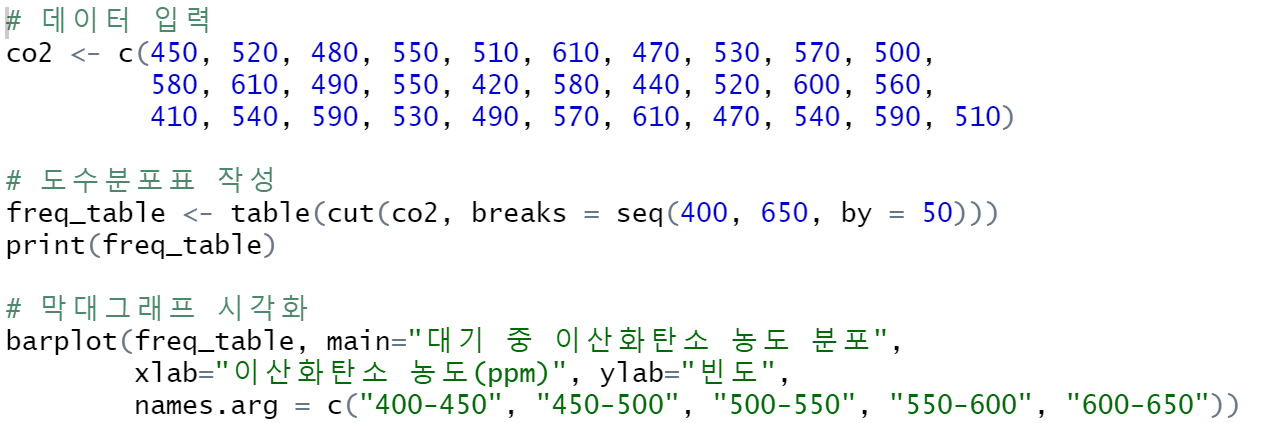
math\_scores <- c(85, 92, 78, 90, 68, 88, 95, 72, 80, 75, 88, 92, 85, 90, 83,

79, 82, 87, 90, 93, 72, 85, 88, 78, 85, 90, 92, 80, 88, 85) # 데이터 할당

**문제2.** [3강. 데이터 수치요약]

어느 도시의 대기오염 정도를 파악하기 위해 한 달 동안 매일 오후 2시에 대기 중 이산화탄소 농도를 측정하였다. 아래의 표는 8월 한 달간 측정된 이산화탄소 농도(ppm)를 나타낸 자료이다. R 프로그램을 사용하여 이 데이터에 대해 도수분포표를 작성하고, 막대그래프로 시각화하시오. 또한 그래프를 통해 알 수 있는 특성을 서술하시오. (25점)





\* 참고: 데이터 입력의 편의를 위해 데이터는 아래와 같이 텍스트로 제공

# 데이터 입력

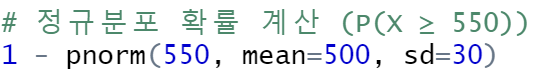
co2 <- c(450, 520, 480, 550, 510, 610, 470, 530, 570, 500,

580, 610, 490, 550, 420, 580, 440, 520, 600, 560,

410, 540, 590, 530, 490, 570, 610, 470, 540, 590, 510)

**문제3.** [6강. 확률분포]

어떤 공장에서 생산되는 제품의 무게는 평균이 500g이고 표준편차가 30g인 정규분포를 따른다고 한다. 임의로 추출된 제품의 무게가 550g 이상일 확률을 R 프로그램을 사용하여 구하고 풀이 과정을 기술하시오. (25점)



**문제4.** [7강. 표본분포]

한 도시의 유권자 중 특정 정책을 지지하는 비율이 70%라고 알려져 있다. 이 정책에 대한 지지율을 추정하기 위해 다양한 크기의 표본을 반복적으로 추출하여 표본비율의 분포를 살펴보고자 한다.

R 프로그램을 사용하여 표본의 크기를 각각 10, 20, 50, 75, 100, 200명으로 설정하고, 각 표본 크기별로 1,000번씩 반복 추출하여 얻은 표본비율들의 평균과 표준편차를 구하시오. 그리고 표본 크기가 커짐에 따라 표본비율 평균이 모비율에 근접해지고, 표본비율의 변동성이 감소하는 경향을 관찰할 수 있는지 서술하시오. (25점)

