출석수업 과제물(평가결과물) 표지(온라인제출용)

**교과목명 : 데이터정보처리입문**

**학 번 : 202334-153942**

**성 명 : 임준희**

**강 의 실 : 광주/전남 지역대학 호**

**연 락 처 : 010–9620-3540**

(1) 2012년 ~ 2022년 연도별 전국 총출생성비 시계열도표

- 2012년부터 2022년까지의 연도별 전국 총출생성비를 나타낸 그래프이다.

2012년 105.7로 시작하여. 2013년, 2014년, 2015년까지 3년동안105.3의 동일한 수치를 유지 하다가 2016년 105로 -0.3이 감소하였다. 그리고 2017년 가장 큰 수치인106.3으로 +1.3이 증가하였고 2018년 다시 105.4으로 -0.9 감소하였다. 2019년 105.5로 증가하였으나, 2020년에 -0.7이 감소하여 가장 낮은 수치인 104.8을 기록하였다. 그리고 2021년, 105.1로 증가하다가 2022년 다시 감소하여 104.7을 기록하였다.

(2) 2012년 ~ 2022년 연도별 전국 합계출산율 시계열도표

2012년부터 2022년까지의 연도별 전국 합계출산율을 나타낸 그래프이다.

2012년부터 2017년까지는1.297, 1.187, 1.205, 1.239, 1.172, 1.052로 1명대를 유지해왔으나 2018년 1명대를 넘기지 못한0.997명을 기록하여, 2022년까지도 0.78로1명대를 기록하지 못했다. 또한 2015년 1.239명을 기록한 후로 2022년까지 합계출산율이 계속 감소하는 추세이다.

<문제 2[[1]](#footnote-1)>

(1) 평균, 중앙값, 표본분산, 표본표준편차, 변동계수

- 10, 56, 20, 84, 38, 78, 57, 97, 26, 78, 77, 32, 24, 22, 34, 82, 58, 21, 18, 81, 42, 47, 65, 37, 29, 11, 17, 76, 48, 66, 59, 95, 11, 31의 데이터에 대해서 평균, 중앙값, 표본분산, 표본표준편차, 변동계수로 나타낸 것이다.

평균은 mean을 사용하여 47.85294, 중앙값은 median을 사용하여 44.5, 표본분산은 var를 사용하여 703.2201, 표본표준편차는 sd를 사용하여 26.5183, 변동계수는 표본표준편차(sd)를 평균(mean)으로 나눈 값인 0.5514623의 결과가 나왔다.

(2) 줄기-잎 그림, 히스토그램, 상자그림

- 위 자료는10, 56, 20, 84, 38, 78, 57, 97, 26, 78, 77, 32, 24, 22, 34, 82, 58, 21, 18, 81, 42, 47, 65, 37, 29, 11, 17, 76, 48, 66, 59, 95, 11, 31의 데이터에 대하여 줄기-잎 그림, 히스토그램, 상자그림으로 시각화 한 것이다.

위 자료들을 보았을 때 20번대가 가장 많으며, 60번대와90번대가 공동으로 가장. 적다는 것을 알 수 있다. 그리고 줄기-잎과, 히스토그램을 통해 분포도가 서로 같은 숫자를 쉽게 파악할 수 있다. 또한, 상자그림을 통해 최솟값부터 제1 사분위수 까지의 영역보다 제3 사분위수 부터 최댓값 까지의 범위가 더 넓다는 것을 확인 할 수 있으며 중간값 부터 제 3사분위가 가장 높은 분포를 보이고 있다.

1. 1) 임의의 데이터로 문제 2번 중 (1), (2) 모두 동일한10, 56, 20, 84, 38, 78, 57, 97, 26, 78, 77, 32, 24, 22, 34, 82, 58, 21, 18, 81, 42, 47, 65, 37, 29, 11, 17, 76, 48, 66, 59, 95, 11, 31을 작성하여 사용 [↑](#footnote-ref-1)